Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses

François-Xavier Berthet,[1]† Hervé G. Zeller,[2] Marie-Thérèse Drouet,[3] Jean Rauzier,[3]† Jean-Pierre Digoutte[1] and Vincent Deubel[3]

[1] Institut Pasteur, BP 220, Dakar, Senegal

[2] Institut Pasteur de Madagascar, BP 1724, Antananarivo, Madagascar

[3] Unité des Arbovirus et Virus des Fièvres Hémorragiques, Institut Pasteur-Paris, 25 rue du Dr Roux, 75724 Paris cedex 15, France

We compared the sequence of an envelope protein gene fragment from 21 temporally distinct West Nile (WN) virus strains, isolated in nine African countries and in France. Alignment of nucleotide sequences defined two groups of viruses which diverged by up to 29%. The first group of subtypes is composed of nine WN strains from France and Africa. The Austral-Asian Kunjin virus was classified as a WN subtype in this first group. The second group includes 12 WN strains from Africa and Madagascar. Four strains harboured a 12 nucleotide in-frame deletion. The loss of the corresponding four amino acids resulted in the loss of the potential glycosylation site present in several WN strains. The distribution of virus subtypes into two lineages did not correlate with host preference or geographical origin. The isolation of closely related subtypes in distant countries is consistent with WN viruses being disseminated by migrating birds.

West Nile (WN) virus is a flavivirus (family *Flaviviridae*) which shares antigenic properties with other members of the Japanese encephalitis (JE) virus serogroup including JE, St Louis encephalitis (SLE), Murray Valley encephalitis (MVE) and Kunjin viruses (Monath & Heinz, 1996). WN virus is a mosquito-borne pathogen responsible for WN fever in humans (Hayes, 1988). Clinical symptoms associated with WN virus infection are most frequently those of a mild febrile illness, but fatal cases of acute meningoencephalitis and fulminant hepatitis have been reported (Monath & Heinz, 1996). WN virus is endemic in tropical areas, particularly India and Africa, where local proliferation of infected mosquitoes can lead to epizootics (Madagascar in 1982 and Senegal in 1988 and 1990) and sudden epidemics (Israel in 1950 and South Africa in 1974). Recent epidemics with high rates of patient mortality have been reported in Algeria (1995) and in Romania (Anon., 1996; Le Guenno *et al.*, 1996) presenting WN virus as a potential emerging human-pathogenic virus. WN virus can infect a wide range of vertebrate species in nature (Hayes, 1988) and wild birds are believed to play an important role in the WN transmission cycle by disseminating the virus during migration (Hayes, 1988; Monath & Heinz, 1996). Because WN virus has a wide geographical distribution, regional variation was investigated from an early stage. Thus, Hammam *et al.* (1965) found that Indian and African WN virus isolates have different haemagglutination inhibition kinetics. Studies of cDNA/RNA heteroduplex restriction profiles and/or reactivity toward monoclonal antibodies (Besselaar & Blackburn, 1988; Mathiot *et al.*, 1990) have identified several WN virus variants. More recently, Porter *et al.* (1993) sequenced NS3 protein gene fragments of seven African and one Indian strains and defined three categories on the basis of nucleotide sequence similarity.

In this study, we report the classification of 21 WN strains (Table 1) isolated in Africa and France by direct nucleotide sequencing of a gene fragment obtained from genomic RNA by RT–PCR. Cytoplasmic RNA was extracted from virus-infected *Aedes pseudoscutellaris* AP61 cells using a previously described method (Deubel *et al.*, 1993). Nucleic acid sequence analysis was performed on an envelope (E) gene fragment obtained by RT–PCR from positions 1318 to 1645 in the WN virus genome (Wengler *et al.*, 1985). The choice was dictated by the variability of this region as demonstrated by comparative amino acid sequence analysis in the flavivirus group. In particular, four amino acids at positions 154 to 157 in the E protein of Kunjin virus (the closest virus to WN virus in the flavivirus evolutionary tree) are absent from the published sequence of a Nigerian WN virus (Wengler *et al.*, 1985).

Author for correspondence: Vincent Deubel.
Fax +33 1 45 68 87 80. e-mail vdeubel@pasteur.fr

† **Present address:** Unité de Génétique Mycobactérienne, Institut Pasteur-Paris, 25 rue du Dr Roux, 75724 Paris cedex 15, France.

The nucleotide sequence data reported in this paper have been deposited in the GenBank database and assigned accession numbers AF001556–AF001574.

Table 1. Characteristics of the 21 West Nile strains analysed in this study

WN strains were obtained from the WHO collaborating centres for reference and research on arboviruses (Pasteur Institutes in Paris and Dakar). Virus strains were recovered from and/or passaged on either suckling mouse brain (before 1990) or cell culture (after 1990).

Strain	Geographical origin	Year of isolation	Primary source of isolation†
Na1047	Kenya	Unknown	(M) *Aedes albothorax*
EntM63134	Uganda	Unknown	Unknown
MP22	Uganda	Unknown	Unknown
HEg101	Egypt	1951	Human
PaH651	France	1965	Human
ArAlg/Djanet	Algeria	1968	(M) *Culex* sp.
ArB310	Central African Republic	1967	(M) *Culex* sp.
AnB3507	Central African Republic	1972	(B) *Antichromus minutus*
ArB3573	Central African Republic	1972	(M) *Culex tigripes*
HB83P55	Central African Republic	1983	Human
HB6343	Central African Republic	1989	Human
AnMg798	Madagascar	1978	(B) *Coracopsis vasa*
ArMg956	Madagascar	1986	(M) *Culex quinquefasciatus*
ArMg978	Madagascar	1988	(M) *Culex univittatus*
AnD27875	Senegal	1979	(P) *Galago senegalensis*
ArD76104*	Senegal	1990	(M) *Mimomyia lacustris*
ArD76986	Senegal	1990	(M) *Culex poicilipes*
ArD78016*	Senegal	1990	(M) *Aedes vexans*
ArD93548	Senegal	1993	(M) *Culex neavei*
ArA3212	Ivory Coast	1981	(M) *Culex guiarti*

* Sequences of the two strains were confirmed using cDNA products obtained by direct amplification of virus RNA in mosquito lysate pools.
† Mosquito (M), bird (B) or primate (P) species.

Moreover, preliminary studies of WN E protein by Western blotting suggested a variability in E protein glycosylation status similar to that observed among Kunjin isolates (V. Deubel, R. Hall & J. Mackenzie, unpublished results; Adams *et al.*, 1995).

cDNA was synthesized from WN virus RNAs using primer WN240 (5′ GAGGTTCTTCAAACTCCAT 3′) and amplified using primers WN240 and WN132 (5′ GAAAACATCAA-GTATGAGG 3′) as described previously (Deubel *et al.*, 1993). Primers WN132 and WN240 correspond to highly conserved sequences in viruses of the JE virus serogroup. DNA amplicons were purified by ion exchange chromatography and precipitated with 2 vols of isopropanol. Each amplicon (1 pmol) was mixed with 1·0 pmol of primer (WN132 or WN240) and sequenced using the *Taq* Dye Deoxy Terminator cycle sequencing kit (Applied Biosystems). Nucleic acid sequences were obtained on an automated Applied Biosystems 373A sequencer. WN virus sequences were aligned with each other using the multiple sequence alignment software CLUSTAL V (Higgins & Sharp, 1988).

The limited sequencing of nucleotides 436 to 690 in the E gene indicated a generally uniform rate of random nucleotide mutations (data not shown). However, four WN strains, including the Nigerian isolate described by Wengler *et al.* (1985), a strain from Uganda of unknown origin and two strains from Senegal isolated in 1990, showed a deletion of 12 nucleotides (nt 462 to 473). The sequenced regions of three of the strains were 100% identical. Each strain was separately cultured in cells and sequences of the two Senegalese strains were identified in RNA extracted from the original mosquito pool, which was preserved at −70 °C, thus excluding the possibility of laboratory contamination (see Table 1). The maximum nucleotide divergence in the sequenced region was 29%. Nucleic acid changes occurred at 99 positions (32% of the gene fragment); 69% of them were in the third codon and 85% were silent.

To visualize the range of relationships among WN viruses, a phylogram was constructed from the nucleotide sequences obtained in this study using the neighbour-joining method (Saitou & Nei, 1987). Because of its similarity to WN virus species, Kunjin virus was also included. The strains formed two distinct lineages (I and II) of closely related subtypes circulating in large and overlapping geographic areas (Fig. 1). The branches leading to these two lineages showed bootstrap values of 1000 and 954 when 1000 trees were sampled, indicating the robustness of the groupings. The lengths of the branches on the phylogenetic tree were proportional to evolutionary distance. Nucleotide sequences of virus subtypes

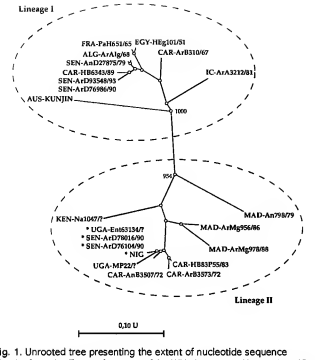

Fig. 1. Unrooted tree presenting the extent of nucleotide sequence identity from the E gene fragment of 21 WN viruses and Kunjin virus (Coia *et al.*, 1988). Nucleotide sequences have been deposited in the GenBank database: SEN-ArD78016, AF001556; CAR-HB83P55, AF001557; CAR-HB6343, AF001558; MAD-AnMg798, AF001559; FRA-PaH651, AF001560; IC-ArA3212, AF001561; UGA-MP22, AF001562; RCA-AnB3507, AF001563; MAD-ArMg956, AF001564; RCA-ArB3573, AF001565; CAR-ArB310, AF001566; ALG-ArDjanet, AF001567; EGY-HEg101, AF001568; SEN-AnD27875, AF001569; SEN-ArD93548, AF001570; KEN-Na1047, AF001571; AUS-Kunjin, AF001572; UGA-Ent63134, AF001573; MAD-ArMg978, AF001574. The geographical origin, strain number and year of isolation are given for each isolate in Table 1. Asterisks indicate the four strains with 12 nucleotide deletions in the E gene fragment. The phylogenetic tree was constructed using the neighbour-joining method (Saitou & Nei, 1987) using the CLUSTAL V program (Higgins & Sharp, 1988). Bootstrap confidence limits between the two defined lineages I and II were calculated from 1000 replicate trees. The horizontal bar indicates the distance corresponding to 10% of nucleotide changes.

in one lineage differed by a maximum of 29% from those of subtypes in the second lineage. Within lineage I the maximum identity of WN strains was 87% and within lineage II it was 80·5%. Kunjin virus shared more than 80% nucleotide identity with WN viruses of lineage I and was consequently classified as a subtype in this lineage. Lineage I is composed of nine WN virus strains from France and North (Algeria, Egypt), West (Senegal, Ivory Coast) and Central (Central African Republic, CAR) Africa. Lineage II includes 12 WN virus strains from West (Senegal, Nigeria), Central (CAR) and East (Uganda, Kenya) Africa and Madagascar. Strains ArD76986 and ArD93548 isolated in Senegal in 1990 and 1993, respectively, have the same E gene fragment sequence. In contrast, strains ArD76104 and ArD78016 isolated in Senegal in 1990 differ from ArD76986, showing that very different subtypes circulate simultaneously in the same country. Three strains from Madagascar were isolated in different regions of the island and from different vectors and hosts. Strain AnMg798, isolated from a parrot, was classified in lineage II with strains ArMg978 and ArMg956 isolated from *Culex* species 8 to 10 years later, but showed more than 19% nucleotide divergence. In lineage II, strains with a 12 nucleotide deletion in the E gene fragment (indicated by an asterisk in Fig. 1) showed less than 5% nucleotide divergence from subtypes from CAR.

Amino acid differences were identified at 15 positions. Alignment of amino acids 146 to 230 of WN virus strains with that of the 1951 Egyptian HEg101 prototype WN strain revealed a maximum divergence of 13% (Fig. 2). Amino acids 154 to 157 were absent from strains ArD76104, ArD78016, EntM63134 and Nigeria. The sequence at these positions in the other strains was NYST, NYPT or SYST. The triad NYS forms an *N*-glycosylation site at position 154 in the E protein in JE, MVE and SLE viruses, and in some Kunjin viruses (Adams *et al.*, 1995). This confirms the previous observation that the E protein of French 1965 WN isolates (V. Deubel, R. Hall & J. Mackenzie, unpublished results) and that of the Sarafend WN strain (Ng *et al.*, 1994) are glycosylated. However, we cannot exclude the possibility that passaging may have influenced the glycosylation status of the E protein of WN strains studied as described for Kunjin virus (Adams *et al.*, 1995). Direct sequencing of cDNA amplified from virus RNA in mosquito pools or in patient sera would confirm the polymorphism of the glycosylation site in natural infecting viruses. Clear signature motifs at amino acid positions A → S 172, N → S 199, T → S 205, T → A 208 and T → S 210 confirmed WN-Kunjin genotype distribution into two lineages. The corresponding published sequence from Kunjin showed only 4 to 11 amino acid changes (5 to 13% divergence) with WN strains, whereas other viruses from the JE serogroup exhibited 24 to 38 amino acid changes (28 to 45% divergence). This result confirmed the close relatedness of Kunjin virus and WN viruses (Coia *et al.*, 1988).

Little information about the genetic diversity of WN viruses was available before this study. Comparative serological analysis of WN strains indicated a variability between isolates from different hosts and vectors and geographic areas but also showed that different strains circulated in the same area at the same time (Gaidamovich & Sokhey, 1973; Hammam *et al.*, 1965; Hammam & Price, 1966; Mathiot *et al.*, 1990; Odelola & Fabiyi, 1976; Price & O'Learly, 1967; Umrigar & Pavri, 1977). We previously demonstrated that the Egyptian prototype HEg101 and the Nigerian strains diverge by 22% in the nucleotide sequences corresponding to the C terminus of NS5 and to part of the 3′ non coding region (Pierre *et al.*, 1994). Porter *et al.* (1993) compared nucleotide sequences at 182 positions in the NS3 gene of eight WN strains from seven countries and reported closer relationships between the Nigerian and Ugandan strains (99·5% similarity) and to a lesser extent with Malagasy strain (86% identity), than with strains from CAR, Ethiopia, Egypt and India which shared more than

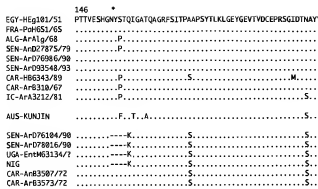

Fig. 2. Amino acid sequences deduced from the 255 nucleotides used to determine genetic relatedness among 21 WN and one Kunjin isolates. Sequences of the Nigerian WN (Wengler *et al.*, 1985), Kunjin (Cola *et al.*, 1988), JE, MVE and SLE (collected in Trent *et al.*, 1987) viruses are aligned. The first amino acid characterized in the WN E protein sequence (Wengler *et al.*, 1985) is indicated. Dots indicate amino acid identities with the Egyptian HEg101 WN and dashes show missing amino acids. The potential *N*-glycosylation site is indicated by an asterisk.

92% identity but diverged by about 24% from the first grouping. Our database on the nucleotide sequence of an E gene fragment confirm this classification and emphasize the substantial genetic variability among African strains, which are unrelated to geographical region, time and host/vector.

Comparative genetic analysis of other flaviviruses including dengue, yellow fever (YF), tick-borne encephalitis (TBE) and JE viruses showed a similar degree of variability within each species, although the classification of TBE and YF viruses correlates with geographic origin and corresponds to topotypes (Zanotto *et al.*, 1996). However, dengue viruses are carried from one continent to the other by viraemic patients travelling by air. Like JE virus, WN virus uses a variety of hosts and vectors for its maintenance and propagation in nature. Wild birds were quickly recognized as important vertebrate hosts for WN virus and their migrations are assumed to be instrumental in virus dissemination (Taylor *et al.*, 1956). Thus, similar subtypes found in France and Senegal and in Senegal and Central Africa suggest continent-to-continent and land-to-land movements. The involvement of birds in the repeated import of WN strains may explain the multiplicity of subtypes in Senegal and in CAR. The data presented in Fig. 1 strongly support the idea that these viruses are free to move between regions and organisms, with no obvious selection.

Sequencing a short fragment of the E gene of WN viruses responsible for recent epidemics, and strains from South Africa and from other parts of the world (former Soviet Union, Israel, Pakistan and India) would give an overall picture of strain variability and virus circulation and may bring new insights into possible relationships with phenotypic variability.

We are grateful to Aïcha Diouf, Annick Leclerc and Mireille Mondo for help in cell culture and virus propagation procedures. We thank Michèle Bouloy and Bahloul Chokri for help in phylogenetic tree analysis and Dr Brigitte Gicquel for providing automated DNA sequencing facilities. François-Xavier Berthet was supported by a doctoral grant from the Agence Nationale de Recherche sur le SIDA (ANRS).

References

Adams, S. C., Broom, A. K., Sammels, L. M., Harnett, A. C., Howard, M. J., Coelen, R. J., Mackenzie, J. S. & Hall, R. A. (1995). Glycosylation and antigenic variation among Kunjin virus isolates. *Virology* **206**, 49–56.

Anon. (1996). Viral meningitidis. *Weekly Epidemiological Record* **71**, 296.

Besselaar, T. G. & Blackburn, N. K. (1988). Antigenic analysis of West Nile virus strains using monoclonal antibodies. *Archives of Virology* **99**, 75–88.

Coia, G., Parker, M. D., Speight, G., Byrne, M. E. & Westaway, E. G. (1988). Nucleotide and complete amino acid sequences of Kunjin virus: definitive gene order and characteristics of the virus-specified proteins. *Journal of General Virology* **69**, 1–21.

Deubel, V., Nogueira, R. M., Drouet, M. T., Zeller, H., Reynes, J. M. & Ha, D. Q. (1993). Direct sequencing of genomic cDNA fragments amplified by the polymerase chain reaction for molecular epidemiology of dengue-2 viruses. *Archives of Virology* **129**, 197–210.

Gaidamovich, S. Y. & Sokhey, J. (1973). Studies on antigenic peculiarities of West Nile virus strains isolated in the USSR by three serological tests. *Acta Virologica* **17**, 343.

Hammam, H. M. & Price, W. H. (1966). Further observations on geographic variation in the antigenic character of West Nile and Japanese B encephalitis viruses. *American Journal of Epidemiology* **83**, 113.

Hammam, H. M., Clark, D. H. & Price, W. H. (1965). Antigenic variations of West Nile virus in relation to geography. *American Journal of Epidemiology* **82**, 40–55.

Hayes, C. G. (1988). West Nile Fever. In *The Arboviruses: Epidemiology and Ecology*, pp. 59–89. Edited by T. P. Monath. Boca Raton, Fla.: CRC Press.

Higgins, D. G. & Sharp, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237–244.

Le Guenno, B., Bougermouth, A., Dif, A., Azzam, T. & Bouakaz, R. (1996). West Nile, a deadly virus? *Lancet* **348**, 1315.

Mathiot, C. C., Georges, A. J. & Deubel, V. (1990). Comparative analysis of West Nile virus strains isolated from human and animal hosts

using monoclonal antibodies and cDNA restriction digest profiles. *Research in Virology* **141**, 533–543.

Monath, T. P. & Heinz, F. X. (1996). Flaviviruses. In *Virology*, pp. 961–1034. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott–Raven.

Ng, M. L., Howe, J., Sreenivasan, V. & Mulders, J. J. L. (1994). Flavivirus West Nile (Sarafend) egress at the plasma membrane. *Archives of Virology* **137**, 303–313.

Odelola, H. A. & Fabiyi, A. (1976). Antigenic relationships among Nigerian strains of West Nile virus by complement fixation and agar gel precipitation techniques. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **70**, 138.

Pierre, V., Drouet, M. T. & Deubel, V. (1994). Identification of mosquito-borne flavivirus sequences using universal primers and reverse transcription/polymerase chain reaction. *Research in Virology* **145**, 93–104.

Porter, K. R., Summers, P. L., Dubois, D., Puri, B., Nelson, W., Henchal, E., Oprandy, J. J. & Hayes, C. G. (1993). Detection of West Nile virus by the polymerase chain reaction and analysis of nucleotide sequence variation. *American Journal of Tropical Medicine and Hygiene* **48**, 440–446.

Price, W. H. & O'Learly, W. (1967). Geographic variation in the antigenic character of West Nile virus. *American Journal of Epidemiology* **85**, 84–86.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular and Biological Evolution* **4**, 406–425.

Taylor, R. M., Work, T. H., Hurlbut, H. S. & Zizk, F. (1956). A study of the ecology of West Nile virus in Egypt. *American Journal of Tropical Medicine and Hygiene* **5**, 579–620.

Trent, D. W., Kinney, R. M., Johnson, B. J., Vorndam, A. V., Grant, J. A., Deubel, V., Rice, C. M. & Hahn, C. S. (1987). Partial nucleotide sequence of St. Louis encephalitis virus RNA: structural proteins, NS1, ns2a and ns2b. *Virology* **256**, 293–304.

Umrigar, M. D. & Pavri, K. K. M. (1977). Comparative serological studies on Indian strains of West Nile virus isolated from different sources. *Indian Journal of Medical Research* **65**, 603.

Wengler, G., Castle, E., Leidner, U., Nowak, T. & Wengler, G. (1985). Sequence analysis of the membrane protein V3 of the flavivirus West Nile virus and of its gene. *Virology* **147**, 264–274.

Zanotto, P. M. D., Gould, E. A., Gao, G. F., Harvey, P. H. & Holmes, E. C. (1996). Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences, USA* **93**, 548–553.

Received 6 March 1997; Accepted 14 May 1997

JOURNAL OF VIROLOGY, June 2002, p. 5847–5856 Vol. 76, No. 12
0022-538X/02/$04.00+0 DOI: 10.1128/JVI.76.12.5847–5856.2002

Infectious cDNA Clone of the Epidemic West Nile Virus from New York City

Pei-Yong Shi,[1,2]* Mark Tilgner,[1] Michael K. Lo,[1] Kim A. Kent,[1] and Kristen A. Bernard[1,2]

Wadsworth Center, New York State Department of Health,[1] and Department of Biomedical Sciences, University at Albany, State University of New York,[2] Albany, New York 12201

Received 15 January 2002/Accepted 18 March 2002

We report the first full-length infectious clone of the current epidemic, lineage I, strain of West Nile virus (WNV). The full-length cDNA was constructed from reverse transcription-PCR products of viral RNA from an isolate collected during the year 2000 outbreak in New York City. It was cloned into plasmid pBR322 under the control of a T7 promoter and stably amplified in *Escherichia coli* HB101. RNA transcribed from the full-length cDNA clone was highly infectious upon transfection into BHK-21 cells, resulting in progeny virus with titers of 1×10^9 to 5×10^9 PFU/ml. The cDNA clone was engineered to contain three silent nucleotide changes to create a *Sty*I site (C to A and A to G at nucleotides [nt] 8859 and 8862, respectively) and to knock out an *Eco*RI site (A to G at nt 8880). These genetic markers were retained in the recovered progeny virus. Deletion of the 3′-terminal 199 nt of the cDNA transcript abolished the infectivity of the RNA. The plaque morphology, in vitro growth characteristics in mammalian and insect cells, and virulence in adult mice were indistinguishable for the parental and recombinant viruses. The stable infectious cDNA clone of the epidemic lineage I strain will provide a valuable experimental system to study the pathogenesis and replication of WNV.

West Nile virus (WNV) is found in many regions, including Africa, the Middle East, Europe, Russia, India, Indonesia, and most recently North America (9). Phylogenetic analysis of WNV strains has revealed two distinct lineages (I and II). Lineage I strains are frequently involved in human and equine outbreaks, and lineage II strains are mostly maintained in enzootic cycles (4, 30, 35, 36, 59). Sequence analysis showed that the strain in North America is closely related to other human epidemic strains isolated from Israel, Romania, Russia, and France, all of which belong to lineage I (35, 36). WNV has caused significant human, equine, and avian disease since its appearance in North America in 1999 (2, 28, 36), and the virus has quickly spread from the Northeast to the eastern seaboard and to the Midwest (3). There were 61 human cases (7 deaths) in New York City in 1999 (13); 21 human cases (4 deaths) in New York, New Jersey, and Connecticut in 2000 (42); and 48 human cases (5 deaths) in New York, Florida, New Jersey, Connecticut, Maryland, Massachusetts, Georgia, and Louisiana in 2001 (14).

WNV is a member of the *Flavivirus* genus, a group of arthropod-borne viruses in the family *Flaviviridae*. Besides WNV, many other members of the flaviviruses are important human pathogens, including dengue virus (DEN), yellow fever virus (YF), the tick-borne encephalitis virus complex (TBE), Japanese encephalitis virus (JE), and Murray Valley encephalitis virus (MVE) (9). The flavivirus genome is a single plus-strand RNA of approximately 11 kb in length that encodes 10 viral proteins in a single open reading frame (55). The encoded polyprotein is translated and co- and posttranslationally processed by viral and cellular proteases into three structural proteins (the capsid protein C; the membrane protein M, which is formed by furin-mediated cleavage of prM; and the envelope protein E) and seven nonstructural proteins (the glycoprotein NS1, NS2a, the protease cofactor NS2b, the protease and helicase NS3, NS4a, NS4b, and the polymerase NS5) (15, 39). The 5′ and 3′ untranslated regions (UTRs) of the genomic RNA are approximately 100 and 400 to 700 nucleotides (nt) in length, respectively, and the terminal nucleotides of both the 5′ and the 3′ UTRs can form highly conserved secondary and tertiary structures (7, 8, 55, 60).

The establishment of a reverse genetic system for the WNV strain presently circulating in the United States is a critical step in the study of the epidemic North American strains of WNV. Infectious full-length cDNA clones for a number of flaviviruses have been successfully developed for the study of viral replication and pathogenesis (56). In several cases, assembly of full-length flavivirus clones in a plasmid vector was not straightforward because clones containing large portions of the genome were unstable and deleterious for bacterial hosts. This problem was first circumvented for YF by ligating cDNA fragments in vitro prior to RNA transcription (54). Similar approaches were applied to develop infectious clones for JE (62), DEN type 2 (DEN2) (31), and TBE strain Hypr (40). For other flaviviruses, stable full-length infectious clones were established for DEN4 (34), Kunjin virus (33), TBE strain Neudoerfl (40), MVE (27), and TBE strain Langat (11). Although an infectious clone of the lineage II WNV strain from Nigeria was recently reported (68), no such full-length cDNA clone has been developed for the human epidemic lineage I WNV.

In this report, we describe the construction of a stable full-length cDNA clone of a WNV strain (lineage I) isolated from the epicenter of New York City during the 2000 outbreak (18). RNA transcribed from the cDNA clone was highly infectious upon transfection into cells, as shown by expression of viral proteins, production of progeny virus, and high specific infectivity. Genetic markers engineered into the cDNA clone to

* Corresponding author. Mailing address: Wadsworth Center, New York State Department of Health, Albany, NY 12201. Phone: (518) 473-7487. Fax: (518) 473-1326. E-mail: ship@wadsworth.org.

distinguish recombinant from parental viruses were retained in the recovered virus. The recombinant virus and the parental virus showed similar biological properties in terms of plaque morphology, growth kinetics, and virulence characteristics. The infectious clone of the present epidemic WNV strain in North America will serve as a valuable reverse genetic system to study the molecular mechanisms of WNV pathogenesis and replication.

MATERIALS AND METHODS

Cells and virus. Vero (ATCC CCL-81) cells were grown in minimal essential medium (MEM) supplemented with 10% fetal bovine serum (FBS). BHK-21/WI2 (BHK-21) (64) and *Aedes albopictus* C6/36 (C6/36) (ATCC CRL-1660) cells were grown in Dulbecco's modification of MEM with 10% FBS and 0.1 mM nonessential amino acids. Antibiotics were added to all media at 10 U/ml of penicillin and 10 μg/ml of streptomycin. Cells were maintained in 5% CO_2 at 37°C (Vero and BHK-21) or 28°C (C6/36). The parental WNV strain 3356 was isolated from the kidney of an American crow collected in October 2000 from Staten Island, New York (18). A single viral stock was made from the second passage in Vero cells without plaque purification, stored as aliquots at −80°C, and designated as parental WNV 3356. This stock was used as parental virus in all assays. Plaque assays were performed on Vero cells as described previously (53).

cDNA synthesis and cloning. BHK-21 cells were infected at a multiplicity of infection (MOI) of 0.05 with parental WNV 3356, and virus was harvested from cell culture media at 36 h postinfection. Genomic RNA was extracted from the cell culture media with RNeasy (Qiagen, Valencia, Calif.). cDNA fragments covering the complete genome were synthesized from genomic RNA through ThermoScript reverse transcription (RT)-PCR according to the manufacturer's instructions (Gibco BRL, Rockville, Md.). Plasmid pBR322 was modified by replacement of the *SphI-EcoRI* fragment in the tetracycline resistance gene with a pair of complementary oligonucleotides to create the sequence 5′-*GCATG*ATCCCGTTGCGCATGCTGATTCGAACCGACTAGT-CTCGAG-TCTAGAATTC-3′ to yield plasmid pBRlinker containing the unique restriction sites *BamH*I, *Sph*I, *Spe*I, *Xho*I, and *Xba*I (listed in order and underlined). After the modification, the original *SphI* site of pBR322 (italics) was mutated through a C to G substitution (lowercase italics). The modified pBR322, pBRlinker, was used as the cloning vector throughout the present experiments.

Bacterial strain HB101 (Gibco BRL) was used as the host for construction and propagation of cDNA clones. Standard cloning procedures were followed (57), except that constructs with inserts of greater than 3 kb were propagated at room temperature. Electroporation was performed to transfect plasmid into bacteria in 0.2-cm cuvettes, using a GenePulser apparatus (Bio-Rad, Hercules, Calif.) with settings of 2.5 kV, 25 μF, and 200 Ω. The virus-specific sequence of each intermediate cloning product was validated by sequence analysis (Applied Biosystems, Foster City, Calif.) before it was used in a subsequent cloning step. All restriction endonucleases were purchased from New England Biolabs (Beverly, Mass.).

RNA transcription and transfection. Plasmid pFLWNV, containing the full-length cDNA of WNV, was amplified in *Escherichia coli* HB101 and purified through MaxiPrep (Qiagen). For in vitro transcription, 5 μg of pFLWNV was linearized with *Xba*I. Mung bean nuclease (5 U; New England BioLabs) was directly added to the *Xba*I digestion reaction mixture, and the reaction mixture was further incubated at 30°C for 30 min to remove the single-stranded nucleotide overhang generated by the *Xba*I digestion. The linearized plasmids were extracted with phenol-chloroform twice, precipitated with ethanol, and resuspended in 10 μl of RNase-free water at 0.5 μg/μl. The mMESSAGE mMACHINE kit (Ambion, Austin, Tex.) was used to in vitro transcribe RNA in a 20-μl reaction mixture with an additional 2 μl of GTP solution. The reaction mixture was incubated at 37°C for 2 h, followed by the addition of DNase I to remove the DNA template. RNA was precipitated with lithium chloride, washed with 70% ethanol, resuspended in RNase-free water, quantitated by spectrophotometry, and stored at −80°C in aliquots. A mutant RNA transcript with a deletion of the 3′-terminal 199 nt of WNV was generated from pFLWNV linearized with an internal restriction site *Dra*I at nt position 10830. The mutant RNA was synthesized in the same manner as the full-length RNA, as described above. All procedures were performed according to manufacturer protocols.

For transfection, approximately 10 μg of RNA was electroporated to 10^7 BHK-21 cells in 0.8 ml of cold phosphate-buffered saline (PBS), pH 7.5, in 0.4-cm cuvettes with the GenePulser apparatus (Bio-Rad) at settings of 0.85 kV and 25 μF, pulsing three times, with no pulse controller. After a 10-min recovery, cells were mixed with media and incubated in a T-75 flask (5% CO_2 at 37°C) until cytopathic effects (CPE) were observed. Virus was harvested as tissue culture media, clarified by centrifugation at 10,000 × *g*, stored in aliquots at −80°C, and designated as recombinant WNV. Plaque assays were performed on Vero cells as described previously (51).

Genetic marker analysis of the recombinant and parental virus. Genetic markers of *Spe*I and *Eco*RI were engineered into the cDNA clone (Fig. 1) to distinguish recombinant progeny virus from the corresponding parental virus. Recombinant virus harvested from supernatant on day 3 posttransfection and parental virus were subjected to RNA extraction with RNeasy (Qiagen). A 388-bp fragment including the genetic markers was amplified through RT-PCR from RNA extracted from either recombinant or parental virus with primers 8706V (5′-CATGGCCATGACTGACACTACTC-3′) and 9093C (5′-CTTGGCCTTTCCGAACTCTCCG-3′). The RT-PCR products were digested with *Sty*I or *Eco*RI and analyzed on a 2% agarose gel.

IFA. Indirect immunofluorescence assays (IFA) were used to detect viral protein expression in WNV RNA-transfected BHK-21 cells. After electroporation, approximately 10^5 transfected cells were spotted onto 10-mm glass coverslips. Cells on coverslips were analyzed by IFA at various times posttransfection for viral protein synthesis. Cells were fixed in 3.7% paraformaldehyde with PBS, pH 7.5, at room temperature for 30 min followed by incubation in −20°C methanol for 30 min. The fixed cells were washed with PBS, incubated at room temperature for 45 min in WNV immune mouse ascites fluid (1:100 dilution; ATCC, Manassas, Va.), and further reacted with goat anti-mouse immunoglobulin G conjugated with fluorescein isothiocyanate at room temperature for 30 min (1:100 dilution) (KPL, Gaithersburg, Md.). The coverslips were washed with PBS, mounted to a slide using fluorescent mounting medium (KPL), and observed under a fluorescence microscope equipped with a video documentation system (Zeiss, Thornwood, N.Y.).

Specific infectivity assay. Approximately 10 μg of RNA was electroporated to 10^7 BHK-21 cells, as described above. Both transfected and untransfected BHK-21 cells were adjusted to a concentration of 6×10^5 cells per ml. A series of 1:10 dilutions were made by mixing 0.5 ml of transfected cells with 4.5 ml of untransfected cells. One milliliter of cells (6×10^4 cells total) for each dilution

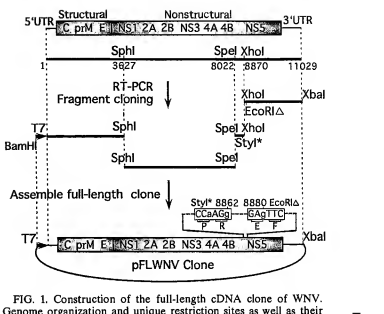

FIG. 1. Construction of the full-length cDNA clone of WNV. Genome organization and unique restriction sites as well as their nucleotide numbers are shown at the top. Four cDNA fragments represented by thick lines were synthesized from genomic RNA through RT-PCR to cover the complete WNV genome. Individual fragments were assembled to form the full-length cDNA clone of WNV (pFLWNV) as described in Materials and Methods. The complete WNV cDNA is positioned under the control of T7 promoter elements for in vitro transcription. Three silent mutations (shown in lowercase) were engineered to create a *Sty*I site (*) and to knock out an *Eco*RI site (Δ) in the NS5 gene. The numbers are the nucleotide positions based on the sequence from GenBank accession no. AF404756.

TABLE 1. Oligonucleotides used to construct the full-length cDNA of WNV

Primer[a]	Primer sequence[b]	Amplified fragment[c]
1V[d]	caaaggatcc*taatacgactcactata*gAGTAGTTCGCCTGTGTGAGCTGA (*Bam*HI)	*Bam*HI-*Sph*I
3839C	ATGTTCTCCTGGTTGGTCCA	
3286V	GTAGAGATTGACTTCGATTAC	*Sph*I-*Spe*I
8804C	CGTACTTCACTCCTTCTGGC	
8016V	GCCCCAACTAGTGCAAAGTTATGGATGGAAC	*Spe*I-*Xho*I
8881C	ATTCTTCTCGAGAGCACATcCTtGGACGTTTTTCTCTGGCC (*Xho*I, *Sty*I)	
8865V	GTGCTCTCGAGAGGAgTTCATAAGA (*Xho*I, *Eco*RIΔ[e])	*Xho*I-*Xba*I
11029C	AacaatctagAGATCCTGTGTTCTCGCACCAC (*Xba*I)	

[a] The primers were named after the nucleotide position of viral sequence and polarity. V, viral genomic sense; C, complementary sense. Nucleotide numbering is based on the sequence from GeneBank accession no. AF404756.

[b] Viral and nonviral sequences are in uppercase and lowercase, respectively. Restriction endonuclease sites are underlined and in parenthesis. Silent mutations within the viral sequences are depicted in lowercase.

[c] cDNA fragments used to construct the full-length clone are shown in Fig. 1.

[d] Italicized sequence represents the T7 promoter.

[e] *Eco*RIΔ represents the knockout of an *Eco*RI site in the cDNA clone by a silent A to g substitution.

was seeded per individual well of six-well plates. Triplicate wells were seeded for each cell dilution. The cells were allowed to attach to the plates for 4 to 5 h under 5% CO_2 at 37°C before the first layer of agar was added, as described previously (53). After incubation of the plates for 3 days under 5% CO_2 at 37°C, a second layer of agar containing neutral red was added. Plaques were counted after incubation of the plates for another 12 to 24 h, and the specific infectivity was calculated as the number of PFU per microgram of RNA.

Growth curves. Subconfluent BHK-21 and C6/36 cells in 12-well plates were inoculated with either the parental or recombinant WNV at an MOI of 5 or 0.05 in triplicate wells. Virus stocks were diluted in BA-1 (M199-H [Gibco-BRL], 0.05 M Tris, pH 7.6, 1% bovine serum albumin, 0.35 g of sodium bicarbonate/liter, 100 U of penicillin/ml, 100 μg of streptomycin/ml, and 1 μg of amphotericin B [Fungizone]/ml). Attachment was allowed for 1 h under 5% CO_2 at 37°C or under 5% CO_2 at 28°C for the BHK-21 and C6/36 cells, respectively. The inocula were then removed, the monolayers were washed three times with BA-1, and 2 ml of medium was added to each well. The plates were incubated for up to 6 days under 5% CO_2 at 37°C or under 5% CO_2 at 28°C for the BHK-21 and C6/36 cells, respectively. The medium was sampled immediately after the addition of medium (1-h time point) and at 7.5, 16, 24, 32, 40, 48 and 72 h for BHK-21 and C6/36 cells, as well as at 96 and 124 h for C6/36 cells. The 10-μl samples were stored at −80°C prior to titration as previously described (53). Cells were observed daily for CPE.

Virulence in mice. Mice were housed in an environmentally controlled room under biosafety level 3 conditions and were given food and water ad libitum. Female outbred CD-1 mice (Charles River Laboratories, Wilmington, Mass.) were obtained at 5 weeks of age and were acclimatized for 1 week. All mice were 6 weeks of age at the start of the experiment. Eight mice per group were inoculated with diluent alone or with 10^2 PFU of parental or recombinant virus subcutaneously (s.c.) in the left rear footpad. Diluent was PBS (endotoxin-free) supplemented with 1% FBS. Mice were evaluated clinically and weighed daily for 2 weeks, then monitored daily and weighed thrice weekly for 2 more weeks. Observed clinical signs included ruffled fur, paresis, hindleg paralysis, and tremors. Morbidity was defined as exhibition of greater than 10% weight loss or clinical signs for 2 or more days. Mice were euthanized if they became moribund. Exposure to virus was confirmed in all surviving mice by a positive antibody titer to WNV by enzyme-linked immunosorbent assay on day 28 postinoculation.

Statistical analyses. Microsoft Excel 97 was used for all statistical analyses. A chi-square test was used to compare the morbidity and mortality in mice for the parental and recombinant viruses, and a two-tailed Student's *t* test was used to evaluate the survival time for the two groups.

RESULTS

Construction and sequencing of the full-length WNV cDNA clone. Selection of the appropriate plasmid vectors and host bacterial strains is critical for construction of flavivirus full-length cDNA clones (56). After testing different vectors and bacterial hosts, plasmid pBR322 and *E. coli* HB101 were chosen for cloning throughout the study. A polylinker containing a number of unique restriction sites was engineered into pBR322 to facilitate the cloning procedure (see Materials and Methods). Cloning procedures for bacterial propagation had higher success rates when performed at room temperature than did those performed at 37°C. All constructs with inserts of greater than 3 kb were propagated at room temperature, but later experiments showed that propagation of the clones at room temperature was not necessarily essential once the clones had been constructed.

Figure 1 shows the overall scheme of the cloning strategy. The primers used to generate individual fragments are listed in Table 1. The full-length cDNA clone was constructed in four steps. First, a fragment from *Spe*I to *Xho*I was amplified by primers 8016V and 8881C and cloned into pBRlinker at their respective sites (Fig. 1), yielding plasmid pSpe-Xho. In order to distinguish the recombinant virus from the parental virus, a *Sty*I site (underlined) was designed in primer 8881C containing silent mutations of nucleotide C to A and A to G at positions 8859 and 8862 (underlined lowercase), respectively. Second, a fragment covering *Xho*I to *Xba*I was amplified by primers 8865V and 11029C and inserted into plasmid pSpe-Xho to generate clone pSpe-Xba (Fig. 1). Primer 8865V contained an A to G substitution at nt 8880 (underlined lowercase) to knock out an *Eco*RI site within the sequence of the parental virus (*Eco*RIΔ in Fig. 1 and Table 1). Third, a fragment covering *Sph*I to *Spe*I was amplified by primers 3286V to 8804C and cloned into pBRlinker at their respective sites, resulting in plasmid pSph-Spe. The fragment from *Sph*I to *Spe*I was then subcloned into pSpe-Xba to yield pSph-Xba (Fig. 1). Finally, a cDNA fragment from *Bam*HI to *Sph*I was amplified by primer 1V and 3839C (Table 1) and cloned into plasmid pSph-Xba at the sites of *Bam*HI and *Sph*I. Primer 1V contained the T7 promoter sequence (italicized lowercase) following the *Bam*HI cloning site (underlined lowercase in Table 1). The resulting plasmid, pFLWNV, contained the complete WNV cDNA under the T7 promoter for in vitro transcription. For RNA synthesis, the full-length cDNA plasmid was linearized by *Xba*I and the overhang nucleotides resulting from *Xba*I digestion were removed by mung bean nuclease. RNA that was in vitro transcribed from this DNA template had an authentic 3′ end of the WNV genomic RNA and a 5′ nonviral G nucleotide derived from the T7 promoter. Analysis of the RNA transcript on a formaldehyde-denaturing agarose gel showed a single band

FIG. 2. Transcription of WNV RNA. Formaldehyde-denaturing 1.0% agarose electrophoresis of RNA transcript together with genomic RNA purified from WNV.

with a mobility identical to that of genomic RNA extracted from WNV (Fig. 2).

Sequence analysis of the full-length cDNA clone showed 11 nucleotide changes compared with the parental virus sequence (35) (Table 2). All nucleotide changes were silent mutations except a T to C transition at nt 7826, which resulted in a conservative change from a valine to an alanine residue. Three nucleotide changes were intentionally designed, two to create a *StyI* site (C to A and A to G at nt 8859 and 8862, respectively) and a third to knock out an *Eco*RI site (A to G at nt 8880). These restriction sites were used as genetic markers to distinguish the recombinant virus from the parental virus (see below). Other mutations in the cDNA clone may derive from the quasispecies of the original virus stock because the parental virus was not plaque purified. It is also possible that some of the mutations occurred during the cloning procedures.

RNA transcript from WNV cDNA clone was highly infectious. Capped RNA transcript was synthesized from the *Xba*I-linearized full-length cDNA plasmid by using T7 RNA polymerase and an optimized 4/3 ratio of methylated cap analogue to GTP. Approximately 30 to 40 μg of RNA was generated from 2 μg of DNA template in a 20-μl reaction mixture. Increasing the ratio of cap analogue to GTP substantially reduced the RNA yield. We routinely electroporated 10 μg of RNA transcript to 10^7 BHK-21 cells as described in Materials and Methods. The transfected cells were incubated and observed for CPE. Apparent CPE were observed in cells on day 3 posttransfection. The cell culture medium was harvested, and virus titer was determined on Vero cells by plaque assays. High viral titers (1×10^9 to 5×10^9 PFU/ml) were consistently obtained.

IFA were used to detect viral protein expression in BHK-21 cells transfected with WNV RNA transcript (Fig. 3). No IFA staining was observed in cells 12 h posttransfection. At 24 h posttransfection, fluorescence was detected in the majority of cells. The staining intensity varied among the IFA-positive cell population. The fluorescent signal increased, and all cells were IFA-positive at 36 h posttransfection. To eliminate the possibility that the positive IFA was merely derived from translation of the transfected RNA rather than from RNA replication in cells, we transfected cells with a mutant RNA containing an expected lethal deletion of the 3′-terminal 199 nt of the genomic RNA. The 3′ deletion RNA was synthesized from the cDNA plasmid digested with a WNV-unique *Dra*I site (nt position 10830). No positive IFA staining was detected in cells at any time points posttransfection (data not shown). These results indicated that the positive IFA signals were initially derived from the replication of RNA in transfected cells and that progeny virus was subsequently generated and spread to neighboring cells through new rounds of infection.

The specific infectivity of RNA transcribed from the full-length cDNA was determined in order to evaluate the efficiency of the system. The specific infectivity of RNA was estimated to be 5×10^4 to 1×10^5 PFU/μg of RNA. Similar specific infectivity was obtained for a mutant RNA containing an extra four nucleotides (5′-CUAG-3′) at the 3′ end. The mutant RNA was synthesized from the *Xba*I-linearized DNA template without mung bean nuclease treatment. Genomic RNA purified from virus showed a specific infectivity of 5×10^5 to 1×10^6 PFU/μg, approximately 10-fold higher than that of transcript RNA. Since uncapped RNA exhibits specific infectivity 10^2- to 10^3-fold lower than that of the capped transcript (40, 54), the discrepancy of infectivity between viral and transcript RNA is most likely due to incomplete capping of the in vitro-transcribed RNA population or to sequence differences, as outlined in Table 2.

The stability of the clone was tested by propagating the *E.coli* HB101 hosting the full-length cDNA plasmid for six continuous passages. Restriction enzyme analysis of the plasmid purified from each of these passages showed digestion patterns identical to that of the original pFLWNV (data not shown). BHK-21 cells transfected with RNA transcripts synthesized from the DNA of passage six showed specific infectivity and CPE indistinguishable from those of passage one cells, indicating that the full-length clone was stable.

Recovered WNV derived from the cDNA clone retained ge-

TABLE 2. Summary of sequence differences between the infectious cDNA clone and parental WNV strain 3356

Nucleotide no.[a]	Strain 3356 genome	cDNA clone	Amino acid change	Location
1285	T	C	Silent	E
3840	T	C	Silent	NS2A
7015	C	T	Silent	NS4B
7826	T	C	V → A	NS5
8067	G	A	Silent	NS5
8859[b]	C	A	Silent	NS5
8862[b]	A	G	Silent	NS5
8880[c]	A	G	Silent	NS5
9123	C	T	Silent	NS5
10613	C	T	Silent	3′UTR
10783	C	T	Silent	3′UTR

[a] Nucleotide position and sequence are based on WNV strain 3356 (GenBank accession no. AF404756).

[b] Mutations were designed to generate an endonuclease *Sty*I site as a marker for recombinant virus (Fig. 4).

[c] This mutation was designed to knock out the endonuclease *Eco*RI site as a marker for recombinant virus (Fig. 4).

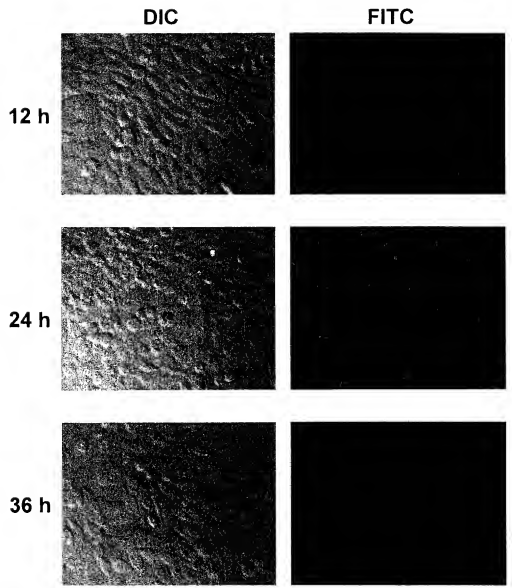

FIG. 3. IFA of viral protein expression in cells transfected with full-length WNV RNA transcript. BHK-21 cells transfected with full-length WNV RNA transcript were analyzed by IFA at the indicated times posttransfection. Photomicrographs were taken at magnifications of ×400. The left and right panels represent the same field of view for each time point. The left panels were visualized with differential interference contrast (DIC), and the right panels were visualized with a fluorescein isothiocyanate (FITC) filter set. For the IFA, WNV immune mouse ascites fluid and goat anti-mouse immunoglobulin G antibody conjugated with FITC were used as primary and secondary antibodies, respectively.

netic markers. To exclude the possibility that the virus recovered from the transfected cells was due to contamination by the parental virus, we engineered genetic markers during the construction of the cDNA clones (Fig. 1). A *Sty*I site was created and an *Eco*RI site was knocked out in the NS5 gene of the recombinant virus. A 388-bp fragment spanning the genetic markers from nt 8706 to 9093 was amplified through RT-PCR from RNA extracted from either parental or recombinant virus. Digestion of the RT-PCR products with *Sty*I and *Eco*RI revealed different cleavage patterns, depending on the origin of the RNA (Fig 4). As expected, PCR products amplified from parental virus were cleaved by *Eco*RI to generate fragments of 173 bp and 215 bp (lane 4 in Fig. 4B), but were not digested by *Sty*I (Fig. 4B, lane 3). In contrast, PCR products derived from the recombinant virus were not cleaved by *Eco*RI (Fig. 4B, lane 8) but were digested by *Sty*I to generate fragments of 152 and 236 bp (Fig. 4B, lane 7). As a negative control, cells were transfected with RNA containing a deletion of the 3′-terminal 199 nt of the WNV genome (Fig. 4B, 3′ dltn). These cells did not yield any RT-PCR product (lane 9). These results clearly show that virus recovered from the transfected cells was derived from the infectious full-length RNA transcript, not from contaminating parental virus.

The phenotypes of recombinant and parental WNV were indistinguishable. There was no difference in plaque size or morphology on Vero cells between the recombinant and pa-

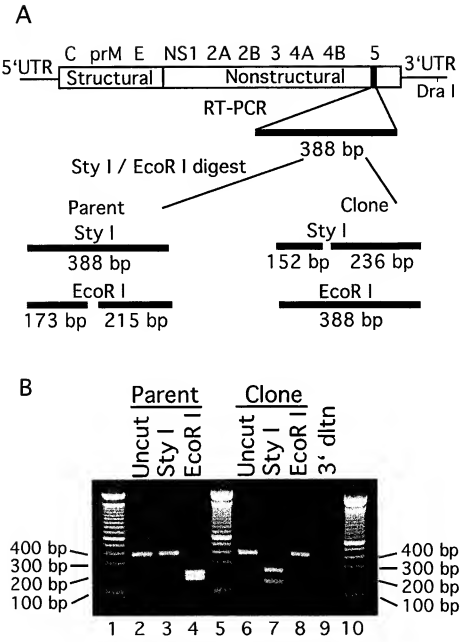

FIG. 4. Recombinant WNV retains the genetic markers engineered during cDNA construction. A *Sty*I site was created and an *Eco*RI site was knocked out in the NS5 gene of the recombinant virus to serve as genetic markers to distinguish recombinant virus from parental virus. A 388-bp fragment (from nt 8706 to 9093) spanning the *Sty*I or *Eco*RI site was amplified through RT-PCR from RNA extracted from either recombinant virus or parental virus. The RT-PCR fragments were subjected to *Sty*I and *Eco*RI digestion. The 388-bp fragment derived from recombinant virus should be cleaved by *Sty*I but not by *Eco*RI; the RT-PCR fragment amplified from parental viral RNA should be digested by *Eco*RI but not *Sty*I. (A) Schematic drawing of genetic marker analysis. The expected sizes of the digestion products are indicated. (B) Agarose gel analysis of genetic markers. Expected digestion pattern as depicted in panel A was observed. As a negative control, no RT-PCR products were detected from the extracted supernatant collected from cells 5 days after transfection with a mutant RNA containing a deletion of the 3'-terminal 199 nt of the genome (lane 9, 3' dltn). A 100-bp ladder was loaded on lanes 1, 5, and 10 as a standard.

rental viruses (Fig. 5). One-step growth curves at a high MOI of 5 were similar for both recombinant and parental viruses on BHK-21 and C6/36 cells, and their growth characteristics at a low MOI of 0.05 were also equivalent on both cell types (Fig. 6). Furthermore, no quantitative or qualitative differences in CPE were observed between the viruses at each MOI. These data suggest that the parental and recombinant viruses are indistinguishable in replication and spread in both mammalian and insect cells.

The virulence phenotypes of the parental and recombinant viruses were compared in adult mice by inoculating 10^2 PFU s.c. The morbidity, mortality, and average survival times are reported in Table 3. The mortality after s.c. inoculation of 10^2 PFU of parental and recombinant viruses was 62 and 50%,

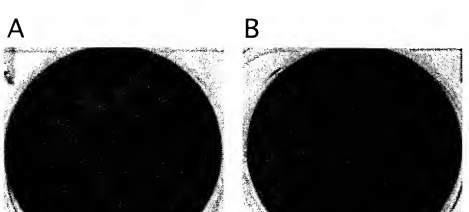

FIG. 5. Plaque morphology of parental and clone-derived WNV on Vero cells. Vero cells in six-well plates were infected with 100 PFU of parental WNV 3356 (A) or 100 PFU of WNV derived from pFLWNV (B). Plaques were visualized 3 days postinoculation by staining for 24 h with neutral red.

respectively. There were no observable differences in the severity or quality of the clinical signs, and the survival curves were very similar. Furthermore, there were no statistical differences in the morbidity ($P = 0.25$), mortality ($P = 0.61$), or average survival times ($P = 0.58$). Thus, the virulence phenotypes were indistinguishable for the parental and recombinant viruses.

DISCUSSION

We report the construction of the first full-length cDNA clone of the human epidemic strain of WNV (lineage I). RNA transcripts transcribed from the cDNA clone were highly infectious upon transfection into BHK-21 cells. The identification of genetic markers engineered into the clone confirmed that the progeny virus was derived from the cDNA clone and thus was not a contaminant. The infectivity of the cDNA-derived RNA was further supported by the finding that a mutant RNA with an expected lethal deletion of the 3′-terminal 199 nt of the genome was not infectious. The recombinant virus showed biological properties indistinguishable from those of the parental virus, including plaque morphology, growth kinetics, and virulence characteristics. These results indicate that an efficient reverse genetic system has been established for lineage I WNV.

A common difficulty in assembling full-length clones of flaviviruses is that plasmids containing long flavivirus-specific inserts are unstable during propagation in bacteria. A number of approaches have been developed to assemble full-length cDNA clones of flaviviruses. (i) Full-length cDNA clone can be assembled through in vitro ligation of cDNA fragments. This approach avoids cloning and propagating full-length clones in bacteria and has been successfully applied to generate infectious RNA of YF (54), JE (62), DEN2 (31), and TBE strain Hypr (40). (ii) Genome-length cDNA containing an upstream promoter for transcription can be directly synthesized by a one-step RT-PCR without any cloning. This rapid method was successfully used to generate infectious RNA of TBE (19). However, this approach has the limitation that the uncloned PCR-derived cDNA will produce a heterogeneous RNA population, derived from mutations during RT-PCR or from quasispecies of the original virus stock. (iii) Full-length cDNA clone can be assembled in yeast cells through homologous recombination. This method was successfully used to assemble a full-length clone of DEN2 (49). (iv) Full-length cDNA can be cloned under the control of a eukaryotic promoter, and introns are introduced into the problematic regions of the cDNA to avoid mutations during their propagation in bacteria. This approach requires transfection of eukaryotic cells with plasmid cDNA rather than RNA. An infectious JE clone was recently developed, using this approach, in which genomic RNA was made in situ by nuclear transcription and intron splicing in transfected eukaryotic cells (67). (v) Stable full-length cDNA clone can be constructed using low- or medium-copy-number vectors and selective bacterial hosts. This approach has been applied to a number of flaviviruses, including DEN4 (34), Kunjin virus (33), TBE strain Neudoerfl (40), MVE (27), TBE strain Langat (38), lineage II WNV (68), and lineage I WNV (this report).

TABLE 3. Morbidity and mortality of parental and recombinant WNV in adult mice[a]

Inoculum	Morbidity (no. sick/total)	Mortality (no. dead/total)	Avg survival time (days [SD])
Diluent	0/8	0/8	NA
Parental WNV	7/8	5/8	9.4 (1.95)
Recombinant WNV	5/8	4/8	10.2 (2.50)

[a] Mice were inoculated with 10^2 PFU s.c. in the left rear footpad. SD, standard deviation; NA, not applicable.

During the construction of our WNV clone, we found that the most unstable region of the genome was within its 5′ quarter and that cDNA from this region should be assembled last in order to obtain full-length clones. These results are consistent with previous reports that cDNAs of structural regions are more likely to be unstable during cloning (67, 68). We also found that bacterial propagation at room temperature

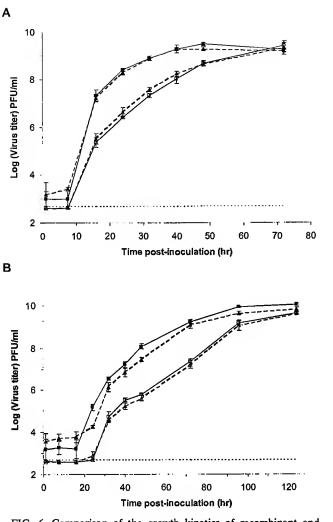

FIG. 6. Comparison of the growth kinetics of recombinant and parental WNV. The growth of recombinant and parental viruses was compared at high and low MOIs on BHK-21 or *Aedes albopictus* C6/36 cells. (A) Growth in BHK-21 cells. (B) Growth in C6/36 cells. Viruses were inoculated at an MOI of 5.0 (filled symbols) or an MOI of 0.05 (open symbols) in triplicate in 12-well plates. Recombinant virus is designated by squares along a solid line. Parental virus is designated by triangles along a dashed line. Error bars represent ± standard deviation of triplicate wells. Dotted line indicates the limit of detection (500 PFU/ml).

rather than at 37°C yielded a higher success rate of cloning intact inserts. A recent report showed that cloning at room temperature was essential to the construction of infectious full-length DEN2 cDNA using high-copy-number plasmid vectors (61).

The 3′ UTR of flaviviruses is believed to function as a promoter for initiation of minus-strand RNA synthesis. The 3′-terminal nucleotides of flavivirus genomic RNA were thermodynamically predicted and experimentally demonstrated to form distinct secondary structures, including a short stem-loop (SL) adjacent to a long SL (8, 50, 52, 60). This secondary structure is conserved among divergent flaviviruses, although only short stretches of sequence are conserved. Structural analysis of WNV 3′ RNA reveals that the loop of the short SL interacts with the lower portion of the neighboring long SL to form a pseudoknot structure (60). Three host proteins bind specifically to the WNV 3′ SL RNA (5), and one of these cellular proteins is the translation elongation factor, eF1-α (6). The NS5 (RNA-dependent RNA polymerase) and NS3 (protease and helicase) of JE were shown to bind specifically to the 3′ SL RNA (17). The cyclization sequences in the 5′- and 3′-terminal regions of the genome were recently demonstrated to be essential for flavivirus replication, both in vivo in Kunjin virus (32) and in vitro in DEN (1, 69). Furthermore, the 3′ SL of WNV RNA was reported to suppress translation of mRNA (38). All of the above reports strongly suggest that the 3′-terminal region of the flavivirus genome plays an essential role in viral replication and possibly regulates translation. Our results show that deletion of the 3′-terminal 199 nt of the genome abolishes the infectivity of WNV RNA. These results agree well with previous reports that deletions in the 3′ UTR of DEN4 (45) and many chimeric 3′ UTRs between DEN2 and WNV (70) are lethal for viral replication. Since the 3′-terminal 199-nt deletion in the mutant RNA included the 3′ cyclization sequence (from nt 10925 to 10932) and the downstream 3′-terminal two-SL structure (from nt 10935 to 11029), it is important to dissect their individual roles during viral replication through systematic mutagenesis analysis.

Many different genetic determinants of virulence have been identified for the flaviviruses (43). For WNV, the studies have been limited. Chambers and coworkers (16) found that neuroinvasion correlates with a mutation in the E gene and determinants outside the E gene. For the related viruses of the JE serocomplex, determinants of neuroinvasion and neurovirulence are in the E gene (12, 21, 22, 37, 44, 47, 65) and the NS1 gene (20). For other flaviviruses, many of the putative virulence determinants are in the E gene (23–26, 29, 48, 58), but mutations in the NS1 gene (10, 46, 51), NS5 gene (66), and the 5′ and 3′ UTR (10, 41) are also associated with virulence. Site-directed mutagenesis of the cDNA clone in this report will allow identification of molecular determinants of virulence for the epidemic strain of WNV.

Lineage I WNV strains have been mostly isolated from epidemic outbreaks and epzootics in birds and equines and have a worldwide distribution. In contrast, lineage II strains have been incidentally isolated from humans with mild febrile disease or without symptoms and are restrictedly found in sub-Saharan Africa and Madagascar (4, 35, 36, 59). Based on sequence analysis of the complete genomes, nucleotide identity between the two lineages is approximately 75% (35). Limited information is known about the pathogenic differences between lineages I and II and among strains within each lineage. Similar growth kinetics were observed for a lineage II Nigerian strain (68) and our lineage I New York strain for both mosquito and mammalian cells. Replacement of the 3′-terminal 1,438 nt of the Nigerian strain (lineage II) with the equivalent sequence (including the complete 3′ UTR and sequence encoding the carboxy-terminal 287 amino acids of NS5) from the prototype WNV Eg101 strain (lineage I) yielded a chimeric virus that showed growth kinetics similar to those of the wild-type Nigerian strain (68). Others have shown differences in neuroinvasiveness in mice between viruses from lineage I and II (D. W. C. Beasley, L. Li, M. T. Suderman, and A. D. Barrett, International Conference on the West Nile Virus, New York Academy of Science Poster Section 1:5, 2001). Lineage I

strains can be further divided into three clades: clade 1a includes viruses from Africa, Europe, Russia, Middle East, and United States; clade 1b includes Kunjin virus from Australia; and clade 1c includes viruses from India (35). Within clade 1a, all U.S. isolates (including the New York crow strain 3356 used in this study) have a nucleotide identity of 99.8%, 99.7% with an Israeli 1998 strain, and 95.2 to 96.4% with strains from Europe, Russia, and Egypt. Kunjin virus in clade 1b exhibits a nucleotide identity of 86.6 to 87.2% with strains in clade 1a (35, 36). During the recent WNV outbreaks, bird mortality was observed in the United States and Israel but not in Europe; therefore, it was speculated that genetic variability within lineage I strains could affect pathogenicity (35, 36, 63). Lanciotti et al. (35) recently showed that six amino acid changes are consistent with the geographic origin of these viruses and might confer the pathogenic difference among these lineage I strains. It will be very interesting to use the infectious clone described here to experimentally test these observations.

Many factors could contribute to the fact that lineage I WNV strains are frequently involved in human outbreaks, while lineage II viruses are mostly maintained in enzootic cycles (4, 36). In addition to possible differences in virulence, differences in vector competence and transmission cycles as well as host immunity may contribute to the difference in disease pattern between the two lineages. The infectious cDNA clones of WNV will serve as a valuable tool to address many of these questions.

ACKNOWLEDGMENTS

We thank Paul Masters for helpful discussions during the course of this work and for critical reading and editing of the manuscript. We thank Margo Brinton for the gift of BHK-21/WI2 cells. We are grateful to many colleagues for technical help. We thank the Molecular Genetics Core at the Wadsworth Center for sequencing and oligonucleotide synthesis.

This work was supported by the New York State Department of Health. M. K. Lo was supported by the Emerging Infectious Diseases Fellowship Program funded by the New York State Department of Health.

REFERENCES

1. **Ackermann, M., and R. Padmanabhan.** 2001. De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. J. Biol. Chem. **276:** 39926–39937.
2. **Anderson, J. F., T. G. Andreadis, C. R. Vossbrinck, S. Tirrell, E. M. Wakem, R. A. French, A. E. Garmendia, and H. J. Van Kruiningen.** 1999. Isolation of West Nile virus from mosquitoes, crows, and a Cooper's hawk in Connecticut. Science **286:**2331–2333.
3. **Bernard, K. A., and L. D. Kramer.** 2001. West Nile virus activity in the United States, 2001. Viral Immunol. **14:**319–338.
4. **Berthet, F. X., H. G. Zeller, M. T. Drouet, J. Rauzier, J. P. Digoutte, and V. Deubel.** 1997. Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. J. Gen. Virol. **78:**2293–2297.
5. **Blackwell, J. L., and M. A. Brinton.** 1995. BHK cell proteins that bind to the 3′ stem-loop structure of the West Nile virus genome RNA. J. Virol. **69:** 5650–5658.
6. **Blackwell, J. L., and M. A. Brinton.** 1997. Translation elongation factor-1 alpha interacts with the 3′ stem-loop region of West Nile virus genomic RNA. J. Virol. **71:**5433–6444.
7. **Brinton, M. A., and J. H. Dispoto.** 1988. Sequence and secondary structure analysis of the 5′-terminal region of flavivirus genome RNA. Virology **162:** 290–299.
8. **Brinton, M. A., A. V. Fernandez, and J. H. Dispoto.** 1986. The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. Virology **153:**113–121.
9. **Burke, D. S., and T. P. Monath.** 2001. Flaviviruses, p. 1043–1126. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, 4th ed., vol. 1. Lippincott William & Wilkins, Philadelphia, Pa.
10. **Butrapet, S., C. Y. Huang, D. J. Pierro, N. Bhamarapravati, D. J. Gubler, and R. M. Kinney.** 2000. Attenuation markers of a candidate dengue type 2 vaccine virus, strain 16681 (PDK-53), are defined by mutations in the 5′ noncoding region and nonstructural proteins 1 and 3. J. Virol. **74:**3011–3019.
11. **Campbell, M. S., and A. G. Pletnev.** 2000. Infectious cDNA clones of Langat tick-borne flavivirus that differ from their parent in peripheral neurovirulence. Virology **269:**225–237.
12. **Cecilia, D., and E. A. Gould.** 1991. Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. Virology **181:**70–77.
13. **Centers for Disease Control and Prevention.** 2000. Guidelines for surveillance, prevention and control of West Nile virus infection—United States. Morb. Mortal. Wkly. Rep. **49:**25–28.
14. **Centers for Disease Control and Prevention.** 2001. Weekly update: West Nile virus activity—United States, November 14–20, 2001. Morb. Mortal. Wkly. Rep. **50:**1061–1062.
15. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. **44:**649–688.
16. **Chambers, T. J., M. Halevy, A. Nestorowicz, C. M. Rice, and S. Lustig.** 1998. West Nile virus envelope proteins: nucleotide sequence analysis of strains differing in mouse neuroinvasiveness. J. Gen. Virol. **79:**2375–2378.
17. **Chen, C. J., M. D. Kuo, L. J. Chien, S. L. Hsu, Y. M. Wang, and J. H. Lin.** 1997. RNA-protein interactions: involvement of NS3, NS5, and 3′ noncoding regions of Japanese encephalitis virus genomic RNA. J. Virol. **71:**3466–3473.
18. **Ebel, G. D., A. P. Dupuis II, K. A. Ngo, D. C. Nicholas, E. B. Kauffman, S. A. Jones, D. Young, J. Maffei, P. Y. Shi, K. A. Bernard, and L. D. Kramer.** 2001. Partial genetic characterization of West Nile virus strains, New York State, 2000. Emerg. Infect. Dis. **7:**650–653.
19. **Gritsun, T. S., and E. A. Gould.** 1995. Infectious transcripts of tick-borne encephalitis virus, generated in days by RT-PCR. Virology **214:**611–618.
20. **Hall, R. A., A. A. Khromykh, J. M. Mackenzie, J. H. Scherret, T. I. Khromykh, and J. S. Mackenzie.** 1999. Loss of dimerisation of the nonstructural protein NS1 of Kunjin virus delays viral replication and reduces virulence in mice, but still allows secretion of NS1. Virology **264:**66–75.
21. **Hasegawa, H., M. Yoshida, T. Shiosaka, S. Fujita, and Y. Kobayashi.** 1992. Mutations in the envelope protein of Japanese encephalitis virus affect entry into cultured cells and virulence in mice. Virology **191:**158–165.
22. **Higgs, S., and E. A. Gould.** 1991. Differences in fusogenicity and mouse neurovirulence of Japanese encephalitis viruses. Arch. Virol. **119:**119–133.
23. **Hiramatsu, K., M. Tadano, R. Men, and C. J. Lai.** 1996. Mutational analysis of a neutralization epitope on the dengue type 2 virus (DEN2) envelope protein: monoclonal antibody resistant DEN2/DEN4 chimeras exhibit reduced mouse neurovirulence. Virology **224:**437–445.
24. **Holbrook, M. R., H. Ni, R. E. Shope, and A. D. Barrett.** 2001. Amino acid substitution(s) in the stem-anchor region of Langat virus envelope protein attenuates mouse neurovirulence. Virology **286:**54–61.
25. **Holzmann, H., F. X. Heinz, C. W. Mandl, F. Guirakhoo, and C. Kunz.** 1990. A single amino acid substitution in envelope protein E of tick-borne encephalitis virus leads to attenuation in the mouse model. J. Virol. **64:**5156–5159.
26. **Holzmann, H., K. Stiasny, M. Ecker, C. Kunz, and F. X. Heinz.** 1997. Characterization of monoclonal antibody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. J. Gen. Virol. **78:** 31–37.
27. **Hurrelbrink, R. J., A. Nestorowicz, and P. C. McMinn.** 1999. Characterization of infectious Murray Valley encephalitis virus derived from a stably cloned genome-length cDNA. J. Gen. Virol. **80:**3115–3125.
28. **Jia, X. Y., T. Briese, I. Jordan, A. Rambaut, H. C. Chi, J. S. Mackenzie, R. A. Hall, J. Scherret, and W. I. Lipkin.** 1999. Genetic analysis of West Nile New York 1999 encephalitis virus. Lancet **354:**1971–1972.
29. **Jiang, W. R., A. Lowe, S. Higgs, H. Reisd, and E. A. Gould.** 1993. Single amino acid codon changes detected in louping ill virus antibody-resistant mutants with reduced neurovirulence. J. Gen. Virol. **74:**931–935.
30. **Jordan, I., T. Briese, and W. I. Lipkin.** 2000. Discovery and molecular characterization of West Nile virus NY 1999. Viral Immunol. **13:**435–446.
31. **Kapoor, M., L. Zhang, P. M. Mohan, and R. Padmanabhan.** 1995. Synthesis and characterization of an infectious dengue virus type-2 RNA genome (New Guinea C strain). Gene **162:**175–180.
32. **Khromykh, A. A., H. Meka, K. J. Guyatt, and E. G. Westaway.** 2001. Essential role of cyclization sequences in flavivirus RNA replication. J. Virol. **75:**6719–6728.
33. **Khromykh, A. A., and E. G. Westaway.** 1994. Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. J. Virol. **68:**4580–4588.
34. **Lai, C. J., B. T. Zhao, H. Hori, and M. Bray.** 1991. Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus. Proc Natl. Acad. Sci. USA **88:**5139–5143.
35. **Lanciotti, R., G. D. Ebel, V. Deubel, A. J. Kerst, S. Murri, R. Meyer, M. Bowen, N. McKinney, W. E. Morrill, M. B. Crabtree, L. D. Kramer, and J. T. Roehrig.** Complete genome sequence and phylogenetic analysis of West Nile virus strains isolated from the United States, Europe and the Middle East. Virology, in press.

36. Lanciotti, R. S., J. T. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. E. Volpe, M. B. Crabtree, J. H. Scherret, R. A. Hall, J. S. MacKenzie, C. B. Cropp, B. Panigrahy, E. Ostlund, B. Schmitt, M. Malkinson, C. Banet, J. Weissman, N. Komar, H. M. Savage, W. Stone, T. McNamara, and D. J. Gubler. 1999. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. Science 286: 2333–2337.
37. Lee, E., and M. Lobigs. 2000. Substitutions at the putative receptor-binding site of an encephalitic flavivirus alter virulence and host cell tropism and reveal a role for glycosaminoglycans in entry. J. Virol. 74:8867–8875.
38. Li, W., and M. A. Brinton. 2001. The 3′ stem loop of the West Nile virus genomic RNA can suppress translation of chimeric mRNAs. Virology 287: 49–61.
39. Lindenbach, B. D., and C. M. Rice. 2001. Flaviviridae: the viruses and their replication, p. 991–1042. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, 4th ed., vol. 1. Lippincott Williams & Wilkins, Philadelphia, Pa.
40. Mandl, C. W., M. Ecker, H. Holzmann, C. Kunz, and F. X. Heinz. 1997. Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr. J. Gen. Virol. 78:1049–1057.
41. Mandl, C. W., H. Holzmann, T. Meixner, S. Rauscher, P. F. Stadler, S. L. Allison, and F. X. Heinz. 1998. Spontaneous and engineered deletions in the 3′ noncoding region of tick-borne encephalitis virus: construction of highly attenuated mutants of a flavivirus. J. Virol. 72:2132–2140.
42. Marfin, A. A., L. R. Petersen, M. Eidson, J. Miller, J. Hadler, C. Farello, B. Werner, G. L. Campbell, M. Layton, P. Smith, E. Bresnitz, M. Cartter, J. Scaletta, G. Obiri, M. Bunning, R. C. Craven, J. T. Roehrig, K. G. Julian, S. R. Hinten, D. J. Gubler, and the A. C. S. Group. 2001. Widespread West Nile Virus activity, Eastern United States, 2000. Emerg. Infect. Dis. 7:730–735.
43. McMinn, P. C. 1997. The molecular basis of virulence of the encephalitogenic flaviviruses. J. Gen. Virol. 78:2711–2722.
44. McMinn, P. C., L. Dalgarno, and R. C. Weir. 1996. A comparison of the spread of Murray Valley encephalitis viruses of high or low neuroinvasiveness in the tissues of Swiss mice after peripheral inoculation. Virology 220: 414–423.
45. Men, R., M. Bray, D. Clark, R. M. Chanock, and C. J. Lai. 1996. Dengue type 4 virus mutants containing deletions in the 3′ noncoding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. J. Virol. 70:3930–3937.
46. Muylaert, I. R., T. J. Chambers, R. Galler, and C. M. Rice. 1996. Mutagenesis of the N-linked glycosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. Virology 222:159–168.
47. Ni, H., and A. D. Barrett. 1998. Attenuation of Japanese encephalitis virus by selection of its mouse brain membrane receptor preparation escape variants. Virology 241:30–36.
48. Pletnev, A. G., M. Bray, and C. J. Lai. 1993. Chimeric tick-borne encephalitis and dengue type 4 viruses: effects of mutations on neurovirulence in mice. J. Virol. 67:4956–4963.
49. Polo, S., G. Ketner, R. Levis, and B. Falgout. 1997. Infectious RNA transcripts from full-length dengue virus type 2 cDNA clones made in yeast. J. Virol. 71:5366–5374.
50. Proutski, V., E. A. Gould, and E. C. Holmes. 1997. Secondary structure of the 3′ untranslated region of flaviviruses: similarities and differences. Nucleic Acids Res. 25:1194–1202.
51. Pryor, M. J., R. C. Gualano, B. Lin, A. D. Davidson, and P. J. Wright. 1998. Growth restriction of dengue virus type 2 by site-specific mutagenesis of virus-encoded glycoproteins. J. Gen. Virol. 79:2631–2639.
52. Rauscher, S., C. Flamm, C. W. Mandl, F. X. Heinz, and P. F. Stadler. 1997. Secondary structure of the 3′-noncoding region of flavivirus genomes: comparative analysis of base pairing probabilities. RNA 3:779–791.
53. Reisen, W. K., R. P. Meyer, S. B. Presser, and J. L. Hardy. 1993. Effect of temperature on the transmission of western equine encephalomyelitis and St. Louis encephalitis viruses by Culex tarsalis (Diptera: Culicidae). J. Med. Entomol. 30:151–160.
54. Rice, C. M., A. Grakoui, R. Galler, and T. J. Chambers. 1989. Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. New Biol. 1:285–296.
55. Rice, C. M., E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, and J. H. Strauss. 1985. Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science 229:726–733.
56. Ruggli, N., and C. M. Rice. 1999. Functional cDNA clones of the Flaviviridae: strategies and applications. Adv. Virus Res. 53:183–207.
57. Sambrook, J., and D. W. Russel. 2001. Molecular cloning, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
58. Sanchez, I. J., and B. H. Ruiz. 1996. A single nucleotide change in the E protein gene of dengue virus 2 Mexican strain affects neurovirulence in mice. J. Gen. Virol. 77:2541–2545.
59. Savage, H. M., C. Ceianu, G. Nicolescu, N. Karabatsos, R. Lanciotti, A. Vladimirescu, L. Laiv, A. Ungureanu, C. Romanca, and T. F. Tsai. 1999. Entomologic and avian investigations of an epidemic of West Nile fever in Romania in 1996, with serologic and molecular characterization of a virus isolate from mosquitoes. Am. J. Trop. Med. Hyg. 61:600–611. (Erratum, 62:162, 2000.)
60. Shi, P. Y., M. A. Brinton, J. M. Veal, Y. Y. Zhong, and W. D. Wilson. 1996. Evidence for the existence of a pseudoknot structure at the 3′ terminus of the flavivirus genomic RNA. Biochemistry 35:4222–4230.
61. Sriburi, R., P. Keelapang, T. Duangchinda, S. Pruksakorn, N. Maneekarn, P. Malasit, and N. Sittisombut. 2001. Construction of infectious dengue 2 virus cDNA clones using high copy number plasmid. J. Virol. Methods 92:71–82.
62. Sumiyoshi, H., C. H. Hoke, and D. W. Trent. 1992. Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. J. Virol. 66:5425–5431.
63. Tsai, T. F. 1998. West Nile encephalitis epidemic in southeastern Romania. Lancet 352:767–771.
64. Vaheri, A., W. D. Sedwick, S. Plotkin, and A. R. Maes. 1965. Cytopathic effect of rubella virus in BHK-21 cells and growth to high titers in suspension culture. Virology 27:239–241.
65. Wu, S. C., W. C. Lian, L. C. Hsu, and M. Y. Liau. 1997. Japanese encephalitis virus antigenic variants with characteristic differences in neutralization resistance and mouse virulence. Virus Res. 51:173–181.
66. Xie, H., K. D. Ryman, G. A. Campbell, and A. D. Barrett. 1998. Mutation in NS5 protein attenuates mouse neurovirulence of yellow fever 17D vaccine virus. J. Gen. Virol. 79:1895–1899.
67. Yamshchikov, V., V. Mishin, and F. Cominelli. 2001. A new strategy in design of +RNA virus infectious clones enabling their stable propagation in E. coli. Virology 281:272–280.
68. Yamshchikov, V. F., G. Wengler, A. A. Perelygin, M. A. Brinton, and R. W. Compans. 2001. An infectious clone of the West Nile flavivirus. Virology 281:294–304.
69. You, S., and R. Padmanabhan. 1999. A novel in vitro replication system for dengue virus. Initiation of RNA synthesis at the 3′-end of exogenous viral RNA templates requires 5′- and 3′-terminal complementary sequence motifs of the viral RNA. J. Biol. Chem. 274:33714–33722.
70. Zeng, L., B. Falgout, and L. Markoff. 1998. Identification of specific nucleotide sequences within the conserved 3′-SL in the dengue type 2 virus genome required for replication. J. Virol. 72:7510–7522.

DOI 10.1099/vir.0.18860-0

A stable full-length yellow fever virus cDNA clone and the role of conserved RNA elements in flavivirus replication

Peter J. Bredenbeek,[1] Engbert A. Kooi,[1]† Brett Lindenbach,[2]‡ Nicolette Huijkman,[1] Charles M. Rice[2]‡ and Willy J. M. Spaan[1]

Correspondence
Peter Bredenbeek
p.j.bredenbeek@lumc.nl

[1]Department of Medical Microbiology, Center of Infectious Disease, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands

[2]Department of Molecular Microbiology, Washington University Medical School, St Louis, MO, USA

Yellow fever virus (YF) is the prototype member of the *Flavivirus* genus. Here, we report the successful construction of a full-length infectious cDNA clone of the vaccine strain YF-17D. YF cDNA was cloned into a low-copy-number plasmid backbone and stably maintained in several *E. coli* strains. Transcribed RNAs had a specific infectivity of 10^8–10^8 p.f.u. (μg RNA)$^{-1}$, and the resulting virus exhibited growth kinetics, plaque morphology and proteolytic processing similar to the parental virus in cell culture. This clone was used to analyse the importance of conserved flavivirus RNA sequences and the 3′ stem–loop structure in virus replication. The conserved sequences 5′CS and CS1, as well as the 3′ stem–loop structure, were found to be essential for virus replication in cell culture, whereas the conserved sequence CS2 and the region containing YF-specific repeated sequences were dispensable. This infectious clone will aid future studies on YF replication and pathogenesis, as well as facilitate the development of YF-17D-based recombinant vaccines.

Received 26 September 2002
Accepted 10 January 2003

INTRODUCTION

Yellow fever virus (YF) is the type member of the genus *Flavivirus*, a group of arthropod-borne RNA viruses in the family *Flaviviridae* (Lindenbach & Rice, 2001). The virus is endemic in the central regions of Africa and South America, where it exists primarily as a zoonosis among monkeys. Occasionally, YF is transmitted to human populations, resulting in epidemics with mortality rates of up to 60 %. Effective protection against infection is possible by vaccination with the attenuated YF-17D strain (reviewed in Burke & Monath, 2001).

The YF genome is a positive-stranded RNA molecule of 11·8 kb, with a 5′ cap structure and a non-polyadenylated 3′ terminus (Rice *et al.*, 1985). The RNA encodes a single open reading frame (ORF) flanked by 5′ and 3′ non-translated regions (NTRs), which are 118 and 565 bases in length, respectively. Translation of YF RNA results in the production of a precursor protein that is cleaved by host and viral proteases to produce the mature viral proteins (see Lindenbach & Rice, 2001, for a review). The N-terminal one-third of this polyprotein encompasses the structural proteins (C-prM-E). Proteolytic processing of the remainder of the polyprotein yields the viral non-structural (NS) proteins (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5). Replication of the viral genome occurs in the cytoplasm and is associated with cellular membranes (reviewed by Lindenbach & Rice, 2001).

All the mosquito-borne flaviviruses share conserved RNA sequences and structures (Fig. 1A). Sequence comparison and RNA secondary structure predictions of the 3′-NTR have revealed several short, well-conserved sequences and indicated that the 3′-terminal region (approximately 90 bases) can be folded in a conserved stem–loop structure (3′-SS) (Brinton *et al.*, 1986; Hahn *et al.*, 1987; Proutski *et al.*, 1997a; Wengler & Castle, 1986). Apart from the sequence 5′-CACAG-3′ in the bulge at the top and A–U and G–C base pairs at the very bottom, this stem–loop structure is not well conserved in primary sequence. A short conserved sequence (CS1; ~26 nucleotides) has been identified upstream of 3′-SS. Complementarity between CS1 and a conserved sequence at the 5′ end of the YF ORF (5′-CS) has been proposed to result in a long-range intramolecular RNA interaction (Hahn *et al.*, 1987). Recent experiments suggest that base-pairing between these sequences is essential

†Present address: Institute for Animal Science and Health (ID-DLO), 8200AB Lelystad, The Netherlands.

‡Present address: Center for the Study of Hepatitis C, Laboratory of Virology and Infectious Disease, The Rockefeller University, New York, USA.

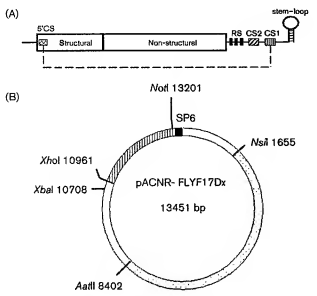

Fig. 1. (A) Schematic representation of the YF genome organization. The large open boxes representing the ORF encoding the viral structural and non-structural proteins is indicated. Cross-hatched box, the conserved sequence (CS) in the capsid gene; black boxes, the repeated sequences; hatched box, CS2; striped box, CS1. The stem–loop structure at the 3′ end is depicted as a hairpin. The conserved structure and sequences are enlarged relative to the size of the ORF. (B) Schematic representation of the pACNR/FLYF-17Dx plasmid. The hatched part is the pACNR vector. The DNA fragments taken from pHYF5′3′IV are indicated by the open boxes and the DNA fragment cloned from pYFM5.2 is indicated by the stippled area. The location of the SP6 RNA polymerase promoter and the restriction enzyme sites that were used to construct this full-length YF cDNA are indicated. *Xho*I was used to linearize the plasmid for *in vitro* RNA transcription.

for RNA replication of a Kunjin virus (KUN) replicon (Khromykh *et al.*, 2001).

The 3′-NTR contains another conserved sequence, CS2, which is approximately 24 nucleotides in length. In YF, CS2 is located 22 nucleotides upstream of CS1. CS2 is duplicated in members of the dengue virus (DEN) and Japanese encephalitis virus (JE) subgroups. In addition to the RNA sequences and structures that are conserved among mosquito-borne flaviviruses, YF has an unique domain consisting of three stretches of closely spaced repeated sequences (RS) that are located just downstream of the YF ORF (Hahn *et al.*, 1987).

Based on computer-assisted RNA folding, phylogenetic sequence comparisons and biochemical and biophysical probing, several models for the RNA structure of the 3′ end of the genome of the mosquito-borne flaviviruses have been proposed (Blackwell & Brinton, 1997; Brinton *et al.*, 1986; Hahn *et al.*, 1987; Olsthoorn & Bol, 2001; Proutski *et al.*, 1999; Shi *et al.*, 1996)). Although these studies do not yield a consensus model for the flavivirus 3′-NTR structure, it is evident that the folding of the 3′-NTR is complex and involves many stem–loop structures and some potential RNA pseudoknots. DEN mutants with deletions in the 3′-UTR have been described (Men *et al.*, 1996), but their analysis does not favour any particular current RNA structure model.

Attempts to construct a stable, full-length infectious YF cDNA in *E. coli* plasmid and λ phage vectors have been unsuccessful due problems with the genetic stability of the full-length clone in the prokaryotic host (Rice *et al.*, 1989). This problem was circumvented by using two plasmids and an *in vitro* ligation approach to create a full-length YF cDNA that could be used for the *in vitro* transcription of infectious YF RNA (Rice *et al.*, 1989). Although cumbersome, this approach yielded the first functional flavivirus cDNA for *in vitro* transcription of infectious YF RNA.

In this study, we describe the construction and characterization of full-length YF cDNA in a low-copy-number vector that is stable in several different bacterial strains. The *in vitro*-transcribed RNA from this clone was shown to be highly infectious. The infectious clone was used to analyse the requirement for the conserved flavivirus RNA elements in YF replication.

METHODS

Cell cultures, virus stocks and plaque assays. BHK-21/J (Lindenbach & Rice, 1997) and SW13 cells were grown in Dulbecco's modified Eagle's medium (DMEM; Life Technologies) supplemented with 7·5 % foetal calf serum (FCS). Virus stocks were obtained by harvesting the medium of BHK-21/J cells that were transfected with *in vitro*-transcribed RNA of wild-type or mutant YF-17D cDNA. The medium was harvested at 36 h post-transfection, clarified by centrifugation (2500 *g* for 10 min) and stored at −80 °C. Virus stocks were titrated by plaque assays on SW13 and BHK21/J cells as described previously (Rice *et al.*, 1989), except that DMEM instead of Eagle's minimal essential medium (MEM) was used in the overlays.

Plasmid constructions. Standard nucleic acid methodologies were used (Ausubel *et al.*, 2000; Sambrook *et al.*, 1989). The *E. coli* strain MC1061 was used for routine cloning purposes, whereas electro-competent *E. coli* Sure cells (Stratagene) were used as a host for the construction of full-length YF cDNAs. The cDNA fragments for the construction of the full-length YF cDNA were taken from the plasmids pHYF5′3′IV and pYFM5.2. Plasmid pYFM5.2 has been described (Rice *et al.*, 1989). pHYF5′3′IV is a phagemid and contains, in addition to the YF sequences of pYF5′3′IV (Rice *et al.*, 1989), the f1 origin for filamentous phage replication. Plasmid pACNR1181 was created from the low-copy-number vector pACNR1180, which contains the polylinker cassette of pSL1180 (Pharmacia) (Ruggli *et al.*, 1996). After digestion of pACNR1180 with *Aat*II and *Not*I, a short spacer sequence was inserted that resulted in the destruction of the *Aat*II site and deleted all the restriction enzyme sites between *Not*I and *Sal*I in the polylinker cassette. The resulting pACNR1181 was used as a vector to assemble a full-length YF-17D cDNA. pHYF5′3′IV was digested with *Not*I and *Xho*I. A 5·1 kb DNA fragment that contained the SP6 RNA polymerase promoter directly fused to the YF 5′ end, the YF 5′ 2271 bp, a small spacer element and the YF 3′ 2586 nucleotides was isolated and cloned in *Not*I/*Xho*I-digested pACNR1181. The resulting plasmid,

pACNR1181YF5'3'IV, was digested with *NsiI* and *AatII* and ligated to the 6747 bp *NsiI–AatII* fragment from pYFM5.2 encompassing the middle part of the YF genome (nt 1655–8402). This resulted in the construction of pACNR/FLYF-17Dx (Fig. 1B), which contained a full-length YF-17D cDNA.

Deletion mutagenesis of conserved RNA sequences and structures. All the deletion mutants were initially created in pHYF5'3'IV. The partial deletion of the YF 5'-CS (Table 1) was constructed by fusion PCR (Landt *et al.*, 1990). All the other deletion mutants were created using uridylated single-stranded pHYF5'3'IV DNA as a template to introduce additional restriction enzyme sites (Kunkel, 1985) flanking the conserved RNA sequence and structural elements in the viral 3'-NTR. *Hind*III sites were inserted immediately 5' and 3' of the RS, CS1 and CS2 sequences. The created plasmids were digested with *Hind*III and religated to create the ΔRS, ΔCS1–CS2, ΔCS1 and ΔCS2 mutants (Table 1). Using the same uridylated template and strategy, an additional *XhoI* site was introduced at nt 10776. This plasmid was cut with *XhoI* and religated to yield the ΔSS mutant (Table 1). The relevant parts of these pHYF5'3'IV derivatives were verified by DNA sequencing and then cloned into the full-length YF cDNA.

***In vitro* transcription.** Plasmids containing the full-length YF cDNA were linearized with *XhoI* and purified by phenol/chloroform extraction and ethanol precipitation. Run-off RNA transcripts were synthesized *in vitro* using SP6 RNA polymerase (Rice *et al.*, 1987). Trace amounts of [^{3}H]UTP were included in the reaction to quantitate the yield. Transcripts were used for transfection without any additional purification.

RNA transfections. A transfection plaque assay (Grakoui *et al.*, 1989) on SW13 cells was used to determine the infectivity of YF-17D transcripts in p.f.u. (μg RNA)$^{-1}$. In short, almost confluent monolayers of SW13 cells in 35 mm diameter dishes were washed twice with PBS lacking Ca^{2+} and Mg^{2+}. A mixture of 0·1–1·0 ng wt or mutant YF-17D RNA transcripts and 4 μg lipofectin (Life Technologies) in 200 μl PBS was added to the cells. After 15 min, the transfection mixture was removed, the cells were washed once with PBS and a DMEM/agarose overlay was applied as described for the plaque assays. Plaques were identified by crystal-violet staining after incubation for 4 days at 37 °C (Rice *et al.*, 1989).

For direct analysis of experiments in which viral RNA synthesis and protein expression was analysed directly in the transfected cells, BHK-21J cells were electroporated with 5 μg of *in vitro*-transcribed YF RNA as described previously (Lindenbach & Rice, 1997). Aliquots were taken from the medium of the transfected cells to quantify the virus yields.

Immunofluorescence. Transfected cells were grown on coverslips. At 24 h post-transfection or infection, the cells were fixed with 3% paraformaldehyde in PBS (pH 7·4) for at least 30 min and washed with PBS containing 10 mM glycine. Following permeabilization with 0·1% Triton X-100 in PBS, the cells were incubated in PBS containing 2% horse serum for 1 h to minimize non-specific immunofluorescence. Indirect immunofluorescence was carried out with a 1:1000 dilution of mAb 1A5 (provided by J.J. Schlesinger) in PBS, which is specific for the YF NS1 protein (Schlesinger *et al.*, 1983) and visualized with a secondary Cy3-conjugated rabbit anti-mouse IgG (Jackson ImmunoResearch Laboratories) diluted 1:1000.

Labelling and analysis of viral RNAs. In general, 3 ml of the electroporated BHK-21J cell suspension (approximately $1·5 \times 10^6$ cells) was placed in 35 mm plates in DMEM containing 2% FCS. At the indicated times post-electroporation (p.e.), the medium was replaced with 750 μl medium containing 2 μg actinomycin D ml^{-1} and 50 μCi [^{3}H]uridine ml^{-1}. At 24 h p.e., RNA was isolated with Trizol (Life Technologies) and resuspended in 21 μl H_2O. One-third of the RNA was denatured with glyoxal and DMSO and analysed by electrophoresis in 0·8% MOPS/agarose gels (Sambrook *et al.*, 1989). Gels were prepared for fluorography as described previously (Bredenbeek *et al.*, 1993).

RESULTS

Construction and characterization of a stable, full-length yellow fever cDNA clone

The low-copy-number vector pACNR1181, derived from pACNR1180 (Ruggli *et al.*, 1996), was used as a vector for the construction of a full-length YF-17D cDNA (see Methods; Fig. 1B). Transformation of *E. coli* Sure cells with the ligation reaction that was expected to yield the full-length YF cDNA construct resulted in the production of small and large bacterial colonies after incubation of the plates for 24 h at 37 °C. Both types of colonies were grown in liquid medium and plasmids were isolated. Based on restriction enzyme digestion patterns all the small colonies ($n=14$) carried the plasmid containing the full-length YF cDNA (pACNR/FLYF-17Dx) whereas the bacteria of the faster-growing

Table 1. Characteristics of YF mutants containing mutations in conserved nucleotide sequences and/or conserved secondary RNA structures

Construct	Position of deleted nucleotides in YF	No. of deleted nucleotides	RNA synthesis*	Plaque formation†
YF-17DΔ5'-CS	155–160	6	−	No
YF-17DΔRS	10333–10520	188 (+aagcuu)‡	++	Yes
YF-17DΔCS1–CS2	10705–10772	68 (+aagcuu)‡	−	No
YF-17DΔCS2	10705–10729	25 (+aagcuu)‡	+	Yes
YF-17DΔCS1	10748–10772	25 (+aagcuu)‡	−	No
YF-17DΔSS	10776–10861	86	−	No
YF-17Dx	−	−	++	Yes

*As determined by [^{3}H]uridine labelling of electroporated BHK-21J cells.

†As determined by lipofection of SW13 cells.

‡Due to engineering an extra *Hind*III site, six extra nucleotides were inserted.

colonies ($n=4$) carried only the parental vector pACNR1181YF5'3'IV (data not shown).

*Xho*I-linearized pACNR/FLYF-17Dx template was used for *in vitro* RNA transcription and the resulting full-length YF transcripts were electroporated into BHK-21J cells. The transfected cells were fixed at 24 h p.e. and analysed for the expression of YF NS1 by immunofluorescence microscopy. Many of the electroporated cells showed a perinuclear, punctated signal when stained with an NS1-specific antibody (Fig. 2A). As a control, cells were transfected with a truncated YF-17Dx transcript lacking the 3' terminal 155 nucleotides of the YF genome (*Xba*I site in Fig. 1B) and therefore unlikely to be replication-competent. The cells transfected with the truncated YF transcripts failed to express any detectable NS1 protein (Fig. 2B). These results demonstrated that YF RNA transcribed from pACNR/FLYF-17Dx was replication-competent and that the NS1 signal was not merely the result of translation of the input RNA.

Stability of the infectious YF cDNA in *E. coli*

The initial *in vitro* ligation approach to obtain infectious YF-17D RNA was developed because of severe stability problems with the full-length YF-17D cDNA in *E. coli*. Therefore it was of great importance to examine the stability of pACNR/FLYF-17Dx in *E. coli*. The recombination-deficient *E. coli* Sure and DH5α strains and recombination-competent *E. coli* MC1061 strain were transformed with pACNR/FLYF-17Dx. Two colonies from each transformation were propagated for ten cycles consisting of alternating growth in liquid medium and plating on a selective medium. No significant change in the growth characteristics or colony morphology of the bacteria was observed during these passages. Restriction enzyme digestion of plasmid DNA purified after 5 and 10 cycles generated the expected DNA pattern (data not shown). YF-17Dx RNA transcribed from these purified plasmids was used to determine the specific infectivity of the YF-17D transcripts. As shown in Table 2, the specific infectivity of the RNA transcripts was comparable [$3{\cdot}5$–$4{\cdot}5 \times 10^5$ plaques (μg RNA)$^{-1}$], irrespective of the *E. coli* host that had been used to grow the plasmid. More importantly, no significant changes in the number of plaques (μg RNA)$^{-1}$ were detected after repeated passages of the plasmid in any of the studied *E. coli* strains.

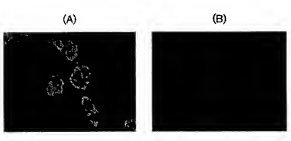

Fig. 2. Immunofluorescence staining of BHK-21 cells with the YF NS1-specific antibody, 1A5, at 24 h p.e. (A) Cells transfected with full-length YF-17Dx transcripts. (B) Cells transfected with YF transcripts lacking the 3' 155 nucleotides of the YF sequence.

Table 2. Specific infectivity of plasmid-derived YF-17Dx RNA after repeated passaging of pANCR-FLYF-17Dx in *E. coli*

E. coli strain	Specific infectivity [YF plaques (μg RNA)$^{-1}$]*		
	Passage 1	Passage 5	Passage 10
Sure	$2{\cdot}8 \times 10^5$	$3{\cdot}8 \times 10^5$	$3{\cdot}5 \times 10^5$
DH5α	ND	$3{\cdot}7 \times 10^5$	$3{\cdot}5 \times 10^5$
MC1061	ND	$4{\cdot}3 \times 10^5$	$4{\cdot}5 \times 10^5$

*For each *E. coli* strain, the specific infectivity is the average of the transfection of *in vitro*-transcribed RNA from two individually passaged bacteria cultures.
ND, Not determined.

Kinetics of viral RNA synthesis and virus production

To study the kinetics of YF RNA synthesis and virus production directly in transfected cells, BHK-21J were electroporated with YF-17Dx transcripts and labelled with [^{3}H]uridine in the presence of actinomycin D for 6 h at 6 h intervals. RNA replication was first detected after labelling the cells from 12 to 18 h p.e. and synthesis peaked between 18 and 30 h (Fig. 3A). After 30 h, viral RNA synthesis decreased, which correlated with the increasing YF-induced cytopathic effect on the transfected cells. Virus could be detected in the medium of the electroporated cells at 12 h p.e. (Fig. 3B). There was a significant rise in virus titre between 12 and 18 h p.e., which levelled off between 24 and 30 h (Fig. 3B) and remained at this level for the remaining period of the experiment.

Next, the growth properties of the parental YF-17D virus and the YF-17Dx virus released into the medium of the transfected cells were compared. BHK-21J cells were infected at an m.o.i. of 10 with a laboratory strain of YF-17D and YF-17Dx virus. Virus was harvested at 12 h intervals and the yield was quantified as described previously (Rice *et al.*, 1989). Similar results were obtained for both virus strains, with virus titres in the medium peaking at 36 h and being slightly lower at later time points (Fig. 4). In addition, the expression of YF proteins was analysed by radioimmunoprecipitation of YF proteins in lysates of [^{35}S] methionine-labelled BHK-21J cells infected with either YF-17D or YF-17Dx virus. No apparent difference in migration or relative levels of the prM, E, NS1, NS2B, NS3, NS4B or NS5 proteins was observed (data not shown).

Conserved flavivirus RNA elements that are essential for virus replication

Several mutant YF-17D cDNAs were constructed by deleting either the conserved flavivirus sequence elements (Δ5'-CS,

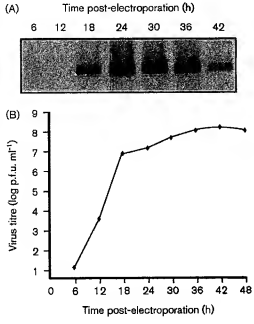

Fig. 3. Kinetics of YF RNA synthesis and virus production in BHK-21J cells electroporated with YF-17Dx transcripts. (A) Kinetics of YF RNA synthesis. The transfected cells were labelled for 6 h with 50 μCi [^{3}H]uridine ml^{-1} in the presence of actinomycin D and lysed at the indicated times p.e. (B) Production of YF in electroporated BHK-21J cells. Samples ($n=2$ per time point) of the medium of the transfected cells were taken at the indicated times p.e. and used for a plaque assay on BHK-21J cells to determine the YF titre.

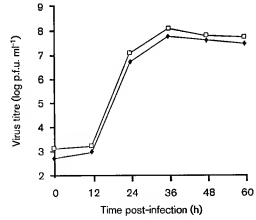

Fig. 4. Comparison of the growth characteristics of YF-17Dx virus derived from the infectious clone (◆) with YF-17D virus stocks derived from vaccine lots (□). BHK-21J cells were infected at an m.o.i. of 10. The medium of the infected cells was sampled at the indicated times post-infection. Plaque assays on BHK-21J cells were used to determine the virus titre.

ΔCS1, ΔCS2 and ΔCS1–CS2), RNA structures (ΔSS), or the subgroup-specific repeated sequences (ΔRS) (Fig. 1A). *In vitro* RNA transcripts were transfected into SW13 and BHK cells by lipofection and electroporation, respectively. Only the ΔRS and ΔCS2 mutants were able to form plaques after lipofection into SW13 cells at either 31 or 37 °C. Plaques were not observed for the Δ5′-CS, ΔCS1–CS2, ΔCS1 and ΔSS mutants. The plaques that were formed by the ΔRS mutant were somewhat smaller than the plaques obtained with YF-17Dx (Fig. 5A). The plaque morphology of the ΔCS2 mutant varied between experiments. Turbid plaques as well as more clear plaques were observed. The size of the ΔCS2 mutant plaques was significantly smaller than observed for YF-17Dx plaques (Fig. 5A).

First cycle RNA analysis was performed to exclude the possibility that the lack of plaque formation for 5′-CS, ΔCS1–CS2, ΔCS1 and ΔSS mutants was due to the deletion of viral sequences that are involved in virion assembly, e.g. the encapsidation signal. BHK-21J cells were electroporated with *in vitro* transcripts of either wt YF-17Dx or the deletion mutants and labelled with [^{3}H]uridine from 18 to 24 h p.e. in the presence of actinomycin D. Total RNA was isolated from the transfected cells and analysed on a denaturing agarose gel. As shown in Fig. 5(B), synthesis of YF RNA could only be detected in the cells transfected with the ΔRS and ΔCS2 mutants. No viral RNA synthesis was detected for the Δ5′-CS, ΔCS1-CS2, ΔCS1 and ΔSS mutants.

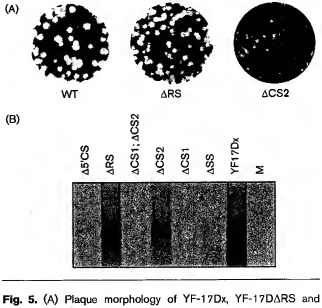

Fig. 5. (A) Plaque morphology of YF-17Dx, YF-17DΔRS and YF-17DΔCS2 after transfection plaque assays on SW13 cells. The transfected cells were incubated at 37 °C, 5% CO_2 for 96 h, fixed with 8% formaldehyde and stained with crystal violet. (B) Analysis YF RNA synthesis in BHK-21J cells electroporated with *in vitro*-transcribed RNA of the indicated YF mutants. Transfected cells were labelled with [^{3}H]uridine in the presence of actinomycin D between 18 and 24 h p.e.

To compare the growth characteristics of the mutants and the parental viruses BHK-21J cells were infected at an m.o.i. of 10 with either the ΔRS or ΔCS2 mutants or with wt YF-17Dx virus. Virus was harvested at 8 h intervals and the yield was determined by plaque assay (Fig. 6). All three viruses showed a rapid increase in titre between 8 and 24 h, with virus production peaking at around 32 h and than levelling off. The kinetics of ΔRS and ΔCS2 virus production appeared a little slower than wt virus and the maximum titre was also somewhat lower. Both the ΔRS and ΔCS2 mutants showed a clear cytopathic effect.

DISCUSSION

Recombinant cDNA clones from which full-length infectious RNA can be transcribed are a valuable tool for studying the molecular biology of positive-strand RNA viruses (Boyer & Haenni, 1994). The approach relies on the infectious nature of the genome RNA of these viruses when transfected into permissive host cells. However, the construction of such functional cDNAs for flaviviruses has proven difficult (Ruggli & Rice, 1999, for a review). Often the plasmids containing a full-length cDNA of these viruses are unstable in *E. coli*. For YF, this problem was initially circumvented by using two plasmids and an *in vitro* ligation approach (Rice *et al.*, 1989). This strategy yielded the first functional flavivirus cDNA for *in vitro* transcription of infectious viral RNA. In this report, the construction and characterization of a stable, full-length YF cDNA in an *E. coli* plasmid vector has been described. The YF cDNA fragments that were used to assemble this clone were taken from the plasmids pYF5′3′IV and pYFM5.2, which were previously used to create the full-length cDNA template for the production of infectious YF RNA by *in vitro* ligation (Rice *et al.*, 1989). The low-copy-number plasmid pACNR1180 was used as a vector for the construction of the full-length YF cDNA. This vector is derived from pACYC177 and has previously been used to construct stable, infectious pestivirus cDNAs (Mendez *et al.*, 1998; Ruggli *et al.*, 1996). The RNA transcribed from pACNR/FLYF-17Dx had a specific infectivity in the order of 10^5–10^6 p.f.u. (μg RNA)$^{-1}$ as determined by lipofection. This was significantly higher than reported for the YF RNA obtained from *in vitro*-ligated templates, even when these values are corrected for the percentage of full-length YF RNA in the RNA transcribed from the ligated template (Rice *et al.*, 1989). Similar specific infectivities have only been reported for a few other flavivirus clones, such as tick-borne encephalitis virus (Mandl *et al.*, 1997) and Murray Valley encephalitis virus (Hurrelbrink *et al.*, 1999). The progeny virus obtained from the pACNR/FLYF-17Dx clone showed similar kinetics in proteolytic processing, viral RNA synthesis and growth as the parental virus in cell culture.

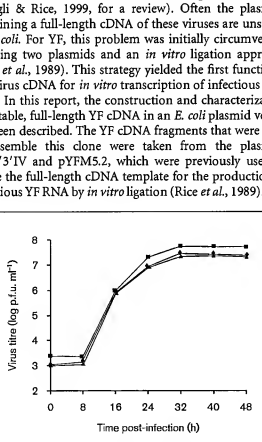

Fig. 6. Growth characteristics YF-17DΔRS (◆) and YF-17DΔCS2 (▲) virus compared with YF-17Dx (■). BHK-21J cells were infected with YF-17Dx, YF-17DΔRS and YF-17DΔCS2 virus at an m.o.i. of 10. Samples of the media were collected at the indicated times post infection. The virus titre was determined by plaque assay on BHK-21J cells. The growth curves are based on the average virus titre in two samples per time point per mutant.

The stability of pACNR/FLYF-17Dx was extensively tested in the *E. coli* strains Sure, DH5α and MC1061. No changes in *E. coli* colony morphology or growth characteristics were observed during the passaging of pACNR/FLYF-17Dx in these bacteria. More importantly, RNA transcribed from plasmid DNA isolated at passages 5 and 10 showed a similar specific infectivity as RNA derived from the originally isolated plasmids. These results demonstrate the successful construction of a stable full-length YF cDNA in a plasmid vector that can be used for *in vitro* transcription of highly infectious viral RNA. Apart from the mutagenesis studies that are described in this report, the pACNR/FLYF-17Dx clone has already been used successfully in other studies (Amberg & Rice, 1999; Kummerer & Rice, 2002; Lindenbach & Rice, 1999).

As for all RNA viruses, the 5′- and 3′-NTRs of the YF genome are believed to play a crucial role in the initiation of viral RNA synthesis. Both the YF 5′- and 3′-NTRs contain sequence motifs and/or stem–loop structures (Fig. 1) that are well conserved among flaviviruses. It has been suggested that these domains are essential for viral RNA synthesis. In this study, the role of the conserved flavivirus sequences (5′-CS, RS, CS2 and CS1) in virus replication was analysed, as well as the stem–loop structure (SS) at the 3′ end of the viral genome.

Deleting the RS domain was well tolerated by YF and resulted in a virus with similar biological properties to the YF-17Dx virus. A domain comparable with the YF RS sequences has also been found in the JE virus serogroup, but there is no sequence similarity in the RS domains of YF and JE-like viruses. For KUN, a deletion of 352 nucleotides in the 5′-proximal half of the 3′-NTR resulted in a five- to tenfold reduction in replication (Khromykh & Westaway, 1997). This deletion was upstream of CS2 and included the RS sequences. RS is lacking in DEN. Deletions in this region of DEN type 4 yielded viruses with delayed growth properties and a smaller plaque size (Men *et al.*, 1996). These deletions

always included the DEN CS2B element and are therefore difficult to compare with the YF-17DΔRS mutant, which retained CS2.

RNA secondary structure analysis of flaviviruses predicts that the YF CS2 sequence forms an independent stem–loop structure within the 3′-NTR (Olsthoorn & Bol, 2001; Proutski *et al.*, 1997b). The fact that this stem–loop structure is well conserved in both pathogenic and vaccine strains of YF suggests that this sequence is essential for virus replication. However, analysis of YF-17DΔCS2 showed that this sequence can be deleted with relatively minor effects on virus replication. Compared with YF-17Dx, the rate of RNA synthesis by YF-17DΔCS2 was somewhat decreased. However, the kinetics of virus production of YFΔCS2 was similar to YF-17Dx. The reason for the observed variability in YF-17DΔCS2 plaque size and morphology is unclear. Small and turbid plaques have also been reported for DEN type 4 deletion mutants involving CS2 (Men *et al.*, 1996). RT-PCR and sequence analysis showed that the deletion was still present in YF-17DΔCS2 progeny virus; however, second-site revertants cannot be excluded.

All the mosquito-borne flaviviruses contain a 5′ conserved sequence that is located a few nucleotides downstream of the translation initiation codon. This 5′-CS is actually part of the flavivirus coding sequence. It has been suggested that the flavivirus 5′-CS sequence base pairs with CS1 via a long-range RNA interaction (Hahn *et al.*, 1987; Khromykh *et al.*, 2001). This interaction is predicted to result in a 'pan-handle'-like structure that is hypothesized to be required for virus replication by modulating virus translation (Khromykh *et al.*, 2001). Recently it was shown, using a KUN replicon and a DEN NS5-based *in vitro* polymerase assay, that complementarity between 5′-CS and CS1 is a prerequisite for viral RNA synthesis (Khromykh *et al.*, 2001; You *et al.*, 2001). The observation that the partial deletion of YF CS1 is lethal for viral RNA synthesis, as reported in this study, is in agreement with a model that requires 'circularization' at some stage of the flavivirus replication cycle. The involvement of CS1 in the 'cyclization' of the viral genomic RNA does not exclude the possibility that the CS1 sequence might also take part in an alternative structure involving base-pairing to other domains within the 3′-NTR. (Proutski *et al.*, 1997a, b; Rauscher *et al.*, 1997; Shi *et al.*, 1996). It can be hypothesized that these different RNA structures involving CS1 are metastable and in equilibrium with each other. This equilibrium may be influenced by factors like RNA–protein interactions and the cellular environment, thereby regulating negative-strand RNA synthesis versus positive-strand RNA synthesis or availability of the RNA for translation.

Another important element of the 3′-NTR RNA is formed by the 3′-terminal 86 nucleotides that are deleted in the YF-17DΔSS mutant. In all the models describing the flavivirus 3′-NTR RNA, these nucleotides are involved in the formation of a hairpin structure (Hahn *et al.*, 1987; Mackenzie *et al.*, 2001; Proutski *et al.*, 1997b; Shi *et al.*, 1996). Deletion of the 3′ 86 nucleotides was lethal for YF RNA synthesis as shown by the results with the YF-17DΔSS mutant.

Finally, as one of the safest and most effective human vaccines, YF-17D recombinants are being vigorously explored as candidate vaccines for other flavivirus diseases, such as JE (Monath, 2002), DEN (Der Most *et al.*, 2000) and West Nile virus (Monath, 2001), as well as for cancer vaccines (McAllister *et al.*, 2000). The availability of full-length stable YF cDNA clones should help in these efforts.

ACKNOWLEDGEMENTS

We thank Linda Lutfiya for initiating the mutagenesis experiments to create the YF deletion mutants. E. K. was supported by Grant 901-02-173 from the Dutch Organization for Sciences (NWO-MW). This work was supported in part by a grant from the Public Health Service (AI31501). We thank Richard Molenkamp for helpful discussions and critical reading of the manuscript.

REFERENCES

Amberg, S. M. & Rice, C. M. (1999). Mutagenesis of the NS2B–NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. *J Virol* **73**, 8083–8094.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (editors) (2000). *Current Protocols in Molecular Biology*. New York: John Wiley.

Blackwell, J. L. & Brinton, M. A. (1997). Translation elongation factor-1 alpha interacts with the 3′ stem–loop region of West Nile virus genomic RNA. *J Virol* **71**, 6433–6444.

Boyer, J. C. & Haenni, A. L. (1994). Infectious transcripts and cDNA clones of RNA viruses. *Virology* **198**, 415–426.

Bredenbeek, P. J., Frolov, I., Rice, C. M. & Schlesinger, S. (1993). Sindbis virus expression vectors: packaging of RNA replicons by using defective helper RNAs. *J Virol* **67**, 6439–6446.

Brinton, M. A., Fernandez, A. V. & Dispoto, J. H. (1986). The 3′-nucleotides of flavivirus genomic RNA form a conserved secondary structure. *Virology* **153**, 113–121.

Burke, D. S. & Monath, T. P. (2001). Flaviviruses. In *Fields Virology*, 4th edn, pp. 1043–1125. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Der Most, R. G., Murali-Krishna, K., Ahmed, R. & Strauss, J. H. (2000). Chimeric yellow fever/dengue virus as a candidate dengue vaccine: quantitation of the dengue virus-specific CD8 T-cell response. *J Virol* **74**, 8094–8101.

Grakoui, A., Levis, R., Raju, R., Huang, H. V. & Rice, C. M. (1989). A cis-acting mutation in the Sindbis virus junction region which affects subgenomic RNA synthesis. *J Virol* **63**, 5216–5227.

Hahn, C. S., Hahn, Y. S., Rice, C. M., Lee, E., Dalgarno, L., Strauss, E. G. & Strauss, J. H. (1987). Conserved elements in the 3′ untranslated region of flavivirus RNAs and potential cyclization sequences. *J Mol Biol* **198**, 33–41.

Hurrelbrink, R. J., Nestorowicz, A. & McMinn, P. C. (1999). Characterization of infectious Murray Valley encephalitis virus derived from a stably cloned genome-length cDNA. *J Gen Virol* **80**, 3115–3125.

Khromykh, A. A. & Westaway, E. G. (1997). Subgenomic replicons of the flavivirus Kunjin: construction and applications. *J Virol* **71**, 1497–1505.

Khromykh, A. A., Meka, H., Guyatt, K. J. & Westaway, E. G. (2001). Essential role of cyclization sequences in flavivirus RNA replication. *J Virol* **75**, 6719–6728.

Kummerer, B. M. & Rice, C. M. (2002). Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. *J Virol* 76, 4773–4784.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A* 82, 488–492.

Landt, O., Grunert, H. P. & Hahn, U. (1990). A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* 96, 125–128.

Lindenbach, B. D. & Rice, C. M. (1997). Trans-complementation of yellow fever virus NS1 reveals a role in early RNA replication. *J Virol* 71, 9608–9617.

Lindenbach, B. D. & Rice, C. M. (1999). Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. *J Virol* 73, 4611–4621.

Lindenbach, B. D. & Rice, C. M. (2001). *Flaviviridae*: The viruses and their replication. In *Fields Virology*, 4th edn, pp. 991–1041. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

McAllister, A., Arbetman, A. E., Mandl, S., Pena-Rossi, C. & Andino, R. (2000). Recombinant yellow fever viruses are effective therapeutic vaccines for treatment of murine experimental solid tumors and pulmonary metastases. *J Virol* 74, 9197–9205.

Mackenzie, J. M., Khromykh, A. A. & Westaway, E. G. (2001). Stable expression of noncytopathic Kunjin replicons simulates both ultrastructural and biochemical characteristics observed during replication of Kunjin virus. *Virology* 279, 161–172.

Mandl, C. W., Ecker, M., Holzmann, H., Kunz, C. & Heinz, F. X. (1997). Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr. *J Gen Virol* 78, 1049–1057.

Men, R., Bray, M., Clark, D., Chanock, R. M. & Lai, C. J. (1996). Dengue type 4 virus mutants containing deletions in the 3′ non-coding region of the RNA genome: analysis of growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. *J Virol* 70, 3930–3937.

Mendez, E., Ruggli, N., Collett, M. S. & Rice, C. M. (1998). Infectious bovine viral diarrhea virus (strain NADL) RNA from stable cDNA clones: a cellular insert determines NS3 production and viral cytopathogenicity. *J Virol* 72, 4737–4745.

Monath, T. P. (2001). Prospects for development of a vaccine against the West Nile virus. *Ann N Y Acad Sci* **951**, 1–12.

Monath, T. P. (2002). Japanese encephalitis vaccines: current vaccines and future prospects. *Curr Top Microbiol Immunol* 267, 105–138.

Olsthoorn, R. C. & Bol, J. F. (2001). Sequence comparison and secondary structure analysis of the 3′ noncoding region of flavivirus genomes reveals multiple pseudoknots. *RNA* 7, 1370–1377.

Proutski, V., Gould, E. A. & Holmes, E. C. (1997a). Secondary structure of the 3′ untranslated region of flaviviruses: similarities and differences. *Nucleic Acids Res* 25, 1194–1202.

Proutski, V., Gaunt, M. W., Gould, E. A. & Holmes, E. C. (1997b). Secondary structure of the 3′-untranslated region of yellow fever virus: implications for virulence, attenuation and vaccine development. *J Gen Virol* 78, 1543–1549.

Proutski, V., Gritsun, T. S., Gould, E. A. & Holmes, E. C. (1999). Biological consequences of deletions within the 3′-untranslated region of flaviviruses may be due to rearrangements of RNA secondary structure. *Virus Res* 64, 107–123.

Rauscher, S., Flamm, C., Mandl, C. W., Heinz, F. X. & Stadler, P. F. (1997). Secondary structure of the 3′-noncoding region of flavivirus genomes: comparative analysis of base pairing probabilities. *RNA* 3, 779–791.

Rice, C. M., Lenches, E. M., Eddy, S. R., Shin, S. J., Sheets, R. L. & Strauss, J. H. (1985). Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science* 229, 726–733.

Rice, C. M., Levis, R., Strauss, J. H. & Huang, H. V. (1987). Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J Virol* 61, 3809–3819.

Rice, C. M., Grakoui, A., Galler, R. & Chambers, T. J. (1989). Transcription of infectious yellow fever RNA from full-length cDNA templates produced by in vitro ligation. *Nat New Biol* 1, 285–296.

Ruggli, N. & Rice, C. M. (1999). Functional cDNA clones of the Flaviviridae: strategies and applications. *Adv Virus Res* 53, 183–207.

Ruggli, N., Tratschin, J. D., Mittelholzer, C. & Hofmann, M. A. (1996). Nucleotide sequence of classical swine fever virus strain Alfort/187 and transcription of infectious RNA from stably cloned full-length cDNA. *J Virol* 70, 3478–3487.

Sambrook, J., Fritsch, T. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schlesinger, J. J., Brandriss, M. W. & Monath, T. P. (1983). Monoclonal antibodies distinguish between wild and vaccine strains of yellow fever virus by neutralization, hemagglutination inhibition, and immune precipitation of the virus envelope protein. *Virology* 125, 8–17.

Shi, P. Y., Brinton, M. A., Veal, J. M., Zhong, Y. Y. & Wilson, W. D. (1996). Evidence for the existence of a pseudoknot structure at the 3′ terminus of the flavivirus genomic RNA. *Biochemistry* 35, 4222–4230.

Wengler, G. & Castle, E. (1986). Analysis of structural properties which possibly are characteristic for the 3′-terminal sequence of the genome RNA of flaviviruses. *J Gen Virol* 67, 1183–1188.

You, S., Falgout, B., Markoff, L. & Padmanabhan, R. (2001). *In vitro* RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5′- and 3′-terminal regions that influence RNA structure. *J Biol Chem* 276, 15581–15591.